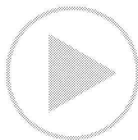
Guide de politique sur le soutien automatisé à la prise de décision

version de 2021



Assurer l'utilisation responsable de l'automatisation, de l'analytique avancée et de l'intelligence artificielle dans la prise de décisions administratives

PRÉAMBULE

Pourquoi avons-nous besoin d'une orientation stratégique dans ce domaine?

À Immigration, Réfugiés et Citoyenneté Canada (IRCC), nous connaissons déjà la prémisse entourant l'automatisation de certains aspects de la prise de décisions administratives : le volume de demandes de résidence temporaire, de résidence permanente et de citoyenneté est en croissance, nos réseaux opérationnels doivent continuellement accroître leur capacité et les clients s'attendent de plus en plus à avoir des échanges numériques rapides et simples. IRCC est engagé dans un processus de transformation profonde et de nouvelles façons d'utiliser la technologie sont un moyen viable de suivre le rythme accéléré tout en maintenant un service de qualité.

Les domaines des données massives et de l'intelligence artificielle (IA) sont en pleine expansion, et IRCC reconnaît désormais leur potentiel pour l'aider à mener ses activités autrement et réaliser des gains d'efficacité. Dans cette nouvelle ère, toutefois, les méthodologies, mécanismes de protection et modèles de gouvernance sont souvent incompatibles avec le fonctionnement des approches algorithmiques. Comme les algorithmes et l'automatisation prennent une place de plus en plus importante, le Ministère bénéficierait d'une approche réfléchie pour résoudre les questions de protection de la vie privée, de gouvernance des données, de sécurité, de transparence, d'équité procédurale et d'interactions humain-machine.

Le gouvernement du Canada a mis en place une nouvelle politique pangouvernementale sur l'utilisation responsable et éthique des systèmes décisionnels automatisés, qui est entrée en vigueur en avril 2020. La *Directive sur la prise de décision automatisée* du Secrétariat du Conseil du Trésor (SCT) soutient la *Politique sur les services et le numérique* et énonce les exigences générales s'appliquant aux ministères utilisant des systèmes, outils et modèles statistiques pour recommander ou prendre des décisions administratives à propos de leurs clients. La Directive est accompagnée de l'outil d'Évaluation de l'incidence algorithmique, qui doit obligatoirement être amorcé dès la phase de conception et finalisé avant qu'un système ne passe d'un environnement d'essai à sa mise en œuvre. Ces instruments stratégiques globaux permettront à IRCC de découvrir ce terrain inconnu.

Pour réussir dans ce domaine, IRCC a aussi besoin d'une orientation claire et facilement accessible qui tient compte à la fois de sa position unique à titre de fournisseur de services pour les Canadiens et les étrangers ainsi que de son contexte opérationnel, ses systèmes et ses programmes particuliers. La politique interne d'IRCC pose des questions fondamentales, notamment : *Dans mon cas, le soutien automatisé à la prise de décision est-il une bonne idée? Avons-nous ce dont nous avons besoin pour réussir? Les avantages surpassent-ils les risques? Comment procéder de façon sécuritaire?* Elle intègre ensuite des concepts d'éthique liés aux technologies dans ses activités quotidiennes et répond directement aux questions que nous nous posons, par exemple : *Pourrions-nous utiliser des algorithmes pour automatiser des décisions défavorables aux demandes de clients? Dans quels cas les « boîtes noires » d'algorithmes sont-elles appropriées, ou ne le sont-elles jamais? Qui est responsable des décisions individuelles prises par une machine? Comment explique-t-on clairement les décisions automatisées aux clients? Comment pouvons-nous nous assurer que les agents d'IRCC profitent des observations tirées d'analyses sans entraver leur prise de décision?*

IRCC doit trouver la réponse à ces questions afin de maintenir la confiance du public, qui est essentielle à la réussite des programmes d'immigration, d'asile, de citoyenneté, d'établissement et de passeport. Si l'on codifie ou « ancre » les règles opérationnelles, que ce soit par l'analytique avancée ou le raisonnement heuristique qu'appliquent depuis longtemps les agents, elles font alors l'objet d'une surveillance accrue. L'application de règles opérationnelles à une grande échelle augmente la probabilité d'incidents, ce qui nécessite des précautions supplémentaires.

En plus d'assurer la gestion des risques prévus, une politique propre au soutien automatisé à la prise de décision peut aider IRCC à réaliser tous les avantages d'une nouvelle approche. Les principes visent à

permettre au personnel d'IRCC de considérer les mesures que devrait éviter le Ministère en raison de leurs risques trop élevés et les activités que nous devrions mener parce qu'elles nous permettent d'offrir de *meilleurs* programmes et services en matière d'immigration.

Quelle est la portée des activités visées par le présent guide?

La présente politique vise **les systèmes automatisés qui appuient, en tout ou en partie, la prise de décisions administratives.** Il s'agit entre autres des systèmes qui : classifient les dossiers selon le niveau de surveillance requis; signalent les dossiers qui devraient faire l'objet d'un examen ou d'une enquête par un agent; formulent des recommandations quant à l'acceptation ou au refus d'une demande; rendent des décisions du début à la fin du processus. Les règles opérationnelles qu'applique aux systèmes automatisés qui soutiennent la prise de décision pourraient découler d'analyses de données sophistiquées ou d'entrevues avec des agents chevronnés; ce que nous savons, c'est que ce système joue aujourd'hui un nouveau rôle dans le modèle de prise de décision d'IRCC. Les règles fondées sur l'analytique sont d'un intérêt particulier pour la présente politique, surtout si l'apprentissage machine entre en ligne de compte.

La politique n'aborde pas tous les types d'automatisation. En effet, elle ne porte pas sur l'automatisation de processus courants, comme la programmation d'ordinateurs pour qu'ils exécutent des tâches de bureau répétitives qui étaient auparavant accomplies manuellement. Elle ne vise pas non plus les systèmes qui aident simplement à la répartition de la charge de travail entre les réseaux, les bureaux et les agents par mesure d'efficacité. Si un processus d'automatisation ne joue pas un certain rôle pour aider IRCC à prendre une décision administrative ayant une incidence sur un client, alors la présente politique n'en tient pas compte. Cependant, on invite le personnel à examiner attentivement si l'automatisation qui semble éloignée de la décision finale pourrait par inadvertance mener à une approbation ou à un refus.

Quelle forme prend le guide?

Pour établir un continuum complet de soutien axé sur l'éthique, il faut à la fois un code d'éthique personnel et un ensemble de principes d'éthique institutionnels. En matière d'éthique personnelle, il n'est pas nécessaire de créer de nouveaux instruments. Le Code de valeurs et d'éthique du secteur public en place s'applique à tous les domaines (il s'applique d'ailleurs déjà tout autant aux comptables menant des vérifications internes qu'aux scientifiques de données qui mettent au point des systèmes d'apprentissage machine). Les cinq valeurs essentielles du Code, soit le respect de la démocratie, le respect envers les personnes, l'intégrité, l'intendance et l'excellence, jettent des bases solides et durables servant à aider les innovateurs d'IRCC à prendre les bonnes décisions pour agir de manière éthique et professionnelle.

En ce qui a trait à l'éthique institutionnelle, un ensemble de principes directeurs est proposé dans la Partie I. **En règle générale, les principes ne dictent pas des mesures précises, mais servent plutôt de fondement pour aider les personnes à prendre des décisions lorsqu'elles sont dans un nouveau milieu et font face à des problèmes inconnus.** En diffusant les principes directeurs présentés dans la Partie I, IRCC espère inspirer, former et outiller ses scientifiques de données, concepteurs de programmes et créateurs de politiques de sorte qu'ils accordent la priorité à l'éthique dans le développement des systèmes de soutien automatisé à la prise de décision. Particulièrement, ils orientent le Ministère dans son ensemble ainsi que les membres de sa direction, qui sont en fin de compte responsables des résultats des systèmes automatisés et de la création d'un milieu de travail qui valorise l'utilisation responsable de la technologie. En tenant compte et en discutant des répercussions générales des efforts novateurs du Ministère, nous ne pouvons que nous assurer que ce dernier appuie de manière constante le bien du public.

Les principes énoncés dans la présente politique sont conçus pour aller de pair avec le guide plus concret contenu dans la Partie II. **Le manuel des responsables de l'automatisation guide le personnel d'IRCC au travers des questions essentielles qui doivent être examinées à différentes étapes d'un projet** : 1) lorsque le soutien automatisé à la prise de décision est envisagé comme une solution possible au problème en cause,

2) lors de la conception d'un nouveau système, 3) lors de la préparation au lancement et 4) une fois qu'un système automatisé est fonctionnel. Le manuel s'adresse à ceux qui prennent part au développement et à la mise en œuvre de systèmes de soutien automatisé à la prise de décision. Il peut notamment s'agir de scientifiques des données, de spécialistes des programmes et des politiques, d'experts en matière de protection de la vie privée, de spécialistes des TI et du personnel opérationnel.

Le guide comprend d'autres chapitres dont **un survol des considérations juridiques et conseils pratiques pour une harmonisation avec le droit administratif, les droits de la personne et la loi sur protection de la vie privée (Partie III)** et **des exigences de base en matière de protection de la vie privée (Partie IV)**. À la fin de la présente politique, on trouve **une liste de vérification visant à favoriser la conformité à la *Directive sur la prise de décision automatisée* (Partie V)** ainsi qu'**un glossaire des termes clés**.

Des questions ou des commentaires au sujet du présent guide? Communiquez avec l'équipe de politiques numériques de la Direction générale des politiques et de la planification stratégiques d'IRCC : IRCC.Digital-Numerique.IRCC@cic.gc.ca.

Une nouvelle politique sur le soutien automatisé à la prise de décision constitue pour le Ministère une occasion de s'assurer que sa réflexion suit le rythme du changement technologique et que ses gens et ses pratiques continuent d'offrir une gamme de programmes à la hauteur des attentes des Canadiens et du monde entier. Les normes et les principes fondamentaux qui guident IRCC demeurent inchangés. Nos engagements, notre mission et nos devoirs restent les mêmes. Même s'il peut sembler que les outils que nous utilisons à l'ère numérique reflètent simplement les pratiques en personne et sur papier d'autrefois, il n'en est rien. L'adoption de nouvelles technologies a changé et changera ce que nous faisons et comment nous le faisons, les compétences qu'il nous faut, nos relations avec nos partenaires et nos clients et notre culture en milieu de travail.

Remarque : Veuillez considérer la présente politique comme un document évolutif qui doit être régulièrement mis à jour pour demeurer pertinent, à mesure qu'évoluent les technologies et nos connaissances à cet égard.

PARTIE I – UN ENSEMBLE DE PRINCIPES DIRECTEURS

Comment le Ministère dans son ensemble devrait-il agir face à des choix difficiles sur certaines applications de technologies émergentes? L'ensemble de principes directeurs qui suit présente les principaux objectifs d'IRCC et couvre divers sujets, tels que les bons outils dans les bonnes circonstances, la conception responsable, la transparence et l'explicabilité, et le recours.

**Cliquez sur les flèches à gauche pour accéder à davantage de renseignements sur chaque principe.*

Qui doit lire cette section?

- *Les cadres supérieurs d'IRCC;*
- *Les praticiens d'IRCC qui œuvrent dans ce domaine;*
- *Toute personne qui s'intéresse à connaître les grands paramètres qui guident IRCC dans son utilisation de l'IA et de l'automatisation.*

1. L'utilisation de l'IA et de l'automatisation devrait s'accompagner d'avantages évidents pour le public. IRCC devrait utiliser ces outils quand il peut le faire de manière responsable, efficace et efficiente, dans cet ordre.

1.1 **En utilisant le soutien automatisé à la prise de décision, IRCC devrait renforcer la confiance du public à l'égard du système d'immigration du Canada**. Le maintien de la confiance des Canadiens est essentiel à la réussite du Ministère et IRCC devrait pallier activement les préoccupations potentielles associées aux données massives et à l'IA. Le Ministère devrait suivre des étapes calculées pour asseoir la légitimité et la confiance du public quant à son utilisation d'outils d'automatisation, d'analytique et d'IA.

1.2 Il existe un coût de renonciation associé au fait de ne pas adopter l'automatisation, l'analytique avancée et l'IA. IRCC ne devrait pas faire preuve de prudence excessive pour ne pas manquer cette occasion au potentiel que représentent les technologies fondées sur les données. Comme dans bien des domaines, **l'utilisation des meilleurs outils à notre disposition fait partie d'un usage efficace et efficient de l'argent des contribuables**. Dans certains cas, les systèmes algorithmiques sont le seul moyen efficient de traiter l'énorme volume d'information requis pour la prestation de services gouvernementaux modernes.

1.3 Cependant, **IRCC peut seulement utiliser ces technologies dans la mesure où leur utilisation est conforme aux normes canadiennes** en matière d'égalité et de non-discrimination, d'équité procédurale, de transparence, de respect de la vie privée, de sécurité et de responsabilisation. L'utilisation responsable de la technologie doit respecter les exigences du droit administratif et des cadres de travail généraux comme la *Charte canadienne des droits et libertés* et la *Loi sur la protection des renseignements personnels*.

1.4 **IRCC ne devrait pas adopter l'approche de décisions automatisées (ou certains aspects des décisions) au détriment de l'intégrité des programmes.** Le travail d'IRCC repose sur la gestion du risque. L'automatisation peut changer les risques ou en introduire de nouveaux, mais l'innovation devrait viser à améliorer l'intégrité des programmes.

1.5 **IRCC devrait déployer tous les efforts possibles pour obtenir des données fiables** pour faire un lien direct entre les résultats des clients et l'information qu'ils soumettent à l'étape de la demande. Simultanément, le **Ministère doit dûment tenir compte de l'impact des conclusions qu'il tire des nouveaux ensembles de données**. Faire la distinction entre la corrélation et la causalité est un défi constant et il est difficile de prévoir avec exactitude les intentions des humains.

1.6 Afin de déterminer des domaines potentiels pour mettre à l'essai l'IA et l'automatisation, IRCC devrait considérer les éléments suivants :

- **Besoin :** Le Canada a connu une croissance année après année du volume de demandes reçues dans plusieurs programmes et on s'attend à ce que cette tendance se poursuive dans un avenir prévisible. Bien que la croissance soit un des principaux moteurs pour l'utilisation de l'IA et de l'automatisation dans les processus décisionnels d'IRCC, l'amélioration de la gestion de l'intégrité des programmes et les attentes de plus en plus élevées des clients sont tout aussi importantes. Le fait de mettre l'accent sur les besoins dès le départ permettra à IRCC de répondre à un problème bien ciblé et d'éviter la mise en place de nouvelles technologies simplement pour « passer au numérique »;
- **Données :** plus la circulation de voyageurs est grande, plus nombreux sont les cas dont on peut tirer des leçons, et là où l'utilisation des demandes en ligne est très répandue, les données sont mieux structurées. La disponibilité de grands volumes de données hautement structurées augmente la qualité des données, ce qui peut réduire des biais involontaires et augmenter l'intégrité des données utilisées pour entraîner des algorithmes et soutenir la prise de décision. Lorsqu'il explorait la possibilité d'utiliser l'apprentissage machine pour soutenir le traitement des demandes de visas de résidents temporaires (VRT), IRCC a examiné un grand nombre de cas où l'intégrité des données était élevée;
- **Enjeux :** la gravité et le caractère définitif de la décision, y compris l'incidence sur les clients, peuvent varier selon le secteur d'activité. De façon générale, on considère que les demandes de VRT comportent des enjeux moindres que les demandes de résidence permanente ou les demandes de citoyenneté, alors que les demandes fondées sur des motifs d'ordre humanitaire ou les cas d'examen des risques avant renvoi comportent des enjeux beaucoup plus grands;
- **Complexité :** les nouvelles technologies sont plus appropriées pour des demandes qui requièrent un examen de routine de l'information, plutôt que des demandes qui font appel à un jugement plus nuancé.

Voir la Directive du SCT, partie 4 (Objectifs et résultats escomptés).

2. *Ce sont des humains (et non des systèmes informatiques) qui sont responsables de la prise de décision, et ce, même lorsque les décisions sont effectuées par des systèmes automatisés.*

2.1 **L'adoption d'un soutien automatisé à la prise de décision peut modifier** ***le moment et l'endroit*** **où l'humain intervient dans le processus de prise de décisions.** Par exemple, l'humain peut jouer un nouveau rôle, comme établir des règles opérationnelles pour un système de triage automatisé à

déployer). **Cependant, les systèmes automatisés ne devraient pas enlever à l'humain son jugement dans la prise de décision**. Les humains jouent un rôle fondamental dans la décision concernant les types de systèmes à utiliser, les dossiers auxquels ces systèmes s'appliquent et les valeurs à programmer dans ces systèmes.

2.2 **Il est important qu'un humain continue d'avoir sa place dans le processus (souvent appelé « comportant l'intervention humaine ») pour inspirer confiance au public. Des modèles qui intègrent un soutien automatisé à la prise de décision devraient prévoir une place pour les humains dans le processus décisionnel.** Des systèmes comportant l'intervention humaine représentent une forme de transparence et de responsabilité qui est mieux connue du public que les processus automatisés. Cela veut dire qu'un décideur humain doit être identifié comme décideur officiel pour toutes les décisions prises ou appuyées par un système décisionnel automatisé.

2.3 Les machines ne peuvent pas se voir transférer les responsabilités lorsqu'elles exécutent des tâches auparavant accomplies par des humains. **IRCC doit assumer la responsabilité des réussites et des échecs de ses systèmes.**

2.4 IRCC devrait aussi réfléchir à la meilleure façon de **consacrer du temps aux questions touchant l'utilisation responsable de l'IA** au sein de ses structures de planification, consultatives et décisionnelles.

Voir la Directive du SCT, parties 6.3.9-6.3.10 (Assurer une intervention humaine).

3. Puisque les décisions d'IRCC ont une incidence importante sur la vie de ses clients et des Canadiens, le Ministère doit accorder la priorité aux approches qui comportent le risque le plus faible.

3.1 L'automatisation des décisions définitives constitue l'approche la plus risquée pour réaliser des gains d'efficacité. IRCC devrait d'abord **examiner s'il peut réaliser des gains d'efficacité par l'automatisation d'autres aspects du processus opérationnel** (p. ex. la répartition de la charge de travail, le triage en fonction des risques, la prise de notes, la recherche et les communications).

3.2 Dans le cas d'un modèle prévisionnel pour la prise de décision, on considère que l'automatisation des approbations a habituellement moins d'incidence que l'automatisation des refus. Cependant, il est important de voir la situation de la perspective de chaque demandeur, particulièrement ceux qui font une demande dans un secteur d'activité où le nombre de places est limité. **Dans les situations où la demande est supérieure à l'offre, les risques des décisions favorables sont plus élevés** : cela peut être perçu comme si l'approbation (ou encore l'invitation ou l'acceptation) d'une demande d'un candidat est accordée au détriment de celle d'un autre.

3.3 **Une autre perspective est celle des Canadiens**, qui profitent des avantages économiques, sociaux et culturels de la migration temporaire et permanente, mais qui aussi s'exposent aux risques de sécurité connexes. **De la perspective collective, les approbations automatisées fondées sur la prévision peuvent avoir des conséquences importantes.** Il est compréhensible que les Canadiens s'inquiètent davantage de l'approbation par erreur de la demande à risque que le refus par erreur

de celle d'un candidat de bonne foi. IRCC doit adopter un point de vue équilibré et tenir compte de l'équité de la perspective individuelle et de la perspective collective.

3.4 Comme IRCC ne peut pas totalement éliminer les risques en matière de sécurité pouvant toucher ses systèmes, il devrait **mettre l'accent sur l'intégration de mesures de sécurité dans la conception de ses modèles de l'IA et son infrastructure afin de minimiser l'impact en cas d'infraction**. Le processus d'évaluation et d'autorisation de sécurité obligatoire appuie les innovateurs dans la prise de décisions réfléchies pour renforcer l'intégrité du système, la gestion de renseignements sensibles et la surveillance.

Voir la Directive du SCT, parties 4 (Objectifs et résultats escomptés), 6.1 (Évaluation de l'incidence algorithmique) et 6.3.7 (Sécurité).

4. Les algorithmes « boîtes noires » peuvent être utiles, mais ne peuvent pas être le seul déterminant dans les décisions définitives concernant les demandes de clients.

4.1 **Les véritables outils d'IA « boîtes noires » sont d'une utilité limitée dans la prise de décisions administratives**, car le Ministère ne devrait pas prendre des décisions qu'il ne peut expliquer clairement. Les outils « boîtes noires », comme la reconnaissance faciale, peuvent jouer un rôle de soutien. **Toutefois, même dans ce cas, le degré de confiance et de fiabilité devrait être considérablement élevé; sinon, des humains devraient pouvoir examiner les résultats du système avant la prise de décisions définitives.**

4.2 **Les algorithmes « boîtes noires » devraient être permis aux fins d'optimisation opérationnelle**, par exemple pour déterminer l'emplacement des bureaux ou comment mieux répartir les ressources entre eux.

Voir la Directive du SCT, partie 6.2.3 (Parvenir des explications après les décisions).

5. IRCC doit reconnaître les limites des technologies fondées sur les données et prendre toutes les mesures raisonnables pour réduire au minimum les préjugés involontaires.

5.1 Chaque algorithme d'apprentissage machine fonctionne entièrement dans un monde défini par les données utilisées pour le calibrer; ainsi, **les limites ou les défaillances dans les ensembles de données nuisent aux résultats**, parfois gravement. Au moment de créer des outils pour la prise de décisions concernant les demandes d'immigration, de citoyenneté ou de passeport, **IRCC doit être prudent pour ne pas reproduire et « ancrer » les préjugés historiques** sous le couvert de la neutralité technologique. Il ne suffit pas de comprendre nos données et notre planification sur le plan technique; nous devons assurer une collaboration interdisciplinaire pour veiller à l'équité de manière concrète.

5.2 Des outils qui fonctionnent avec l'IA peuvent aussi servir à étudier des programmes existants d'une nouvelle façon afin de veiller à l'atteinte des résultats escomptés et de **mettre en lumière des préjugés qui auraient pu se glisser involontairement dans les pratiques de longue date** des décideurs humains.

5.3 Tous les outils analytiques sont limités par leur capacité de prévoir et de décrire les résultats. Bien que les tendances générales soient utiles, elles ne peuvent pas remplacer la preuve d'actions personnelles. C'est pourquoi **IRCC doit être très prudent lorsqu'il utilise de l'information sur un groupe ou une personne**.

5.4 **Les données massives devraient aider IRCC à se tourner vers le passé comme vers l'avenir.** IRCC devrait envisager d'utiliser de nouveaux outils pour détecter, évaluer et éliminer tout préjugé ou toute incohérence.

5.5 **IRCC devrait chercher à améliorer continuellement la qualité de ses fonds de données, puisque l'analytique avancée et l'IA reposent sur des données de qualité.** En accordant la priorité au développement d'outils fondés sur les données, nous accordons aussi la priorité à la portée, à la profondeur et à la fiabilité de nos ensembles de données.

5.6 Il est possible que l'analytique laisse supposer qu'une demande représente un faible risque parce que son profil est similaire à celui de candidats approuvés par le passé, et dans ce cas, **IRCC pourrait souhaiter assumer le risque d'une approbation automatisée**. **Cependant, le calcul n'est pas le même dans le scénario inverse** : un refus automatisé ne peut pas être fondé que sur des similarités entre le profil d'un demandeur et celui de candidats refusés par le passé. **Un refus doit être fondé sur des *données probantes* sur la demande en question** Par exemple, un refus automatisé fondé sur le fait que le passeport d'un demandeur de visa de résidence temporaire figure dans la base de données des documents de voyage perdus et volés d'Interpol serait justifiable, car il s'appuie sur un fait vérifiable et objectif sur le demandeur en question (c'est-à-dire qu'il ne respecte pas l'exigence d'un document de voyage valide). À l'inverse, il ne serait pas justifiable qu'un refus automatisé s'appuie sur le fait que le pays d'origine du demandeur affiche un taux élevé de fraude de passeport. Il peut s'agir d'un indicateur de risque parmi d'autres, mais ce n'est pas une raison suffisante en soi pour justifier un refus catégorique.

Voir la Directive du SCT, parties 6.3.1-6.3.2 (Évaluer et surveiller les résultats) et 6.3.3 (Qualité des données).

6. Les agents devraient être informés, et non menés vers une conclusion.

6.1 Les agents peuvent s'appuyer sur les systèmes de soutien automatisé à la prise de décision en usant de leur jugement. Dans ces cas, les systèmes doivent être conçus de manière à ce qu'**ils n'entravent pas les décideurs** dans l'exercice de leurs capacités et pouvoirs.

6.2 **Les agents ont besoin d'une formation et d'un contexte appropriés pour comprendre ce que le système de soutien aux décisions leur indique précisément**. Par exemple, les étiquettes ont une importance : lorsque le système trie automatiquement les demandes des clients en groupes, il est important de savoir si ces groupes portent l'étiquette « vert/jaune/rouge » ou « simple/moyen/complexe ». La première étiquette peut sembler représenter des instructions à l'intention des agents (vert=approuver, rouge=refuser) et la dernière, une simple indication du temps et des efforts requis que devra consacrer l'agent pour prendre une décision.

6.3 Les données massives devraient **renforcer la prise de décisions administratives en fournissant une foule de renseignements pertinents aux décideurs et en orientant leur attention vers les facteurs déterminants**. De cette façon, les nouveaux outils devraient aider les agents à prendre des décisions plus rapidement, mais surtout à prendre de *meilleures* décisions.

6.4 **IRCC devrait s'efforcer de fournir à ses décideurs la meilleure information et les meilleurs outils disponibles.** En mettant à la disposition des décideurs de nouvelles connaissances et en leur permettant de centrer leur attention sur les facteurs les plus importants, nous éliminerons les efforts inutiles, améliorerons l'uniformité et lierons mieux l'action et sa justification.

6.5 Des essais avec des utilisateurs pendant la conception et le développement d'un projet peuvent permettre au Ministère d'**acquérir une meilleure compréhension de l'interaction des humains avec des systèmes automatisés**. Au-delà de fournir des indications sur les façons les plus efficaces et fiables de présenter des extrants générés par des machines aux agents, une compréhension approfondie des dynamiques en jeu peut renforcer la responsabilité et aider à prévenir des conséquences négatives involontaires.

Voir la Directive du SCT, partie 6.3.5 (Formation des employés).

7. Les humains et l'IA remplissent des rôles complémentaires. IRCC doit s'efforcer de préciser les rôles de chacun.

7.1 L'automatisation, l'analytique avancée et l'IA sont des multiplicateurs de puissance; elles peuvent **aider et accroître les capacités des personnes**. L'utilisation de ces technologies afin de permettre au personnel d'IRCC de se concentrer sur l'identification de solutions créatives ou du travail visant les clients devrait améliorer à la fois la prestation de services et la satisfaction professionnelle.

7.2 **IRCC devrait miser sur les forces des ordinateurs**, soit d'analyser avec fiabilité de grands volumes de données et d'examiner des millions de possibilités sans se fatiguer, **et permettre aux êtres humains de miser sur les leurs**, soit leur intuition, leur créativité, leur empathie, leurs compétences sociales et leur capacité de façonner une stratégie globale. **Les systèmes algorithmiques devraient permettre aux humains de se concentrer sur des aspects qui requièrent leur expertise et leur jugement.**

7.3 **Les systèmes qui prennent eux-mêmes les décisions administratives**, au lieu d'appuyer les décideurs humains, **conviennent pour les décisions concernant des déterminations simples et factuelles**, notamment le fait que le client possède un dossier criminel ou est déjà venu au Canada. L'automatisation devrait être centrée d'abord et avant tout sur des tâches routinières, où il est facile de décrire les résultats souhaités et où il n'est pas nécessaire que l'humain exerce son jugement.

7.4 En revanche, lorsque des systèmes automatisés visent à appuyer des décideurs humains, la formation devrait être conçue de façon à inciter les agents à **faire appel à leur propre jugement professionnel et à leur intuition**, afin d'améliorer la prise de décision[1].

---

[1] https://www.nesta.org.uk/blog/human-vs-machine-why-we-need-be-thinking-more-about-how-humans-are-interacting-ai-tools/ (en anglais seulement)

7.5 **La présence d'un humain dans le processus assure non seulement une protection contre les risques à l'équité procédurale, mais elle produit aussi de meilleurs résultats.** Pour les évaluations simples (par exemple la mise en correspondance d'une personne et de sa photo), le rendement des algorithmes fondés sur les données est souvent supérieur à celui de l'humain. Quand toutefois les machines et les humains unissent leurs forces, les résultats sont encore meilleurs. L'une des façons d'améliorer le rendement est de demander à des experts d'appliquer leur jugement dans les extrants de modèles mathématiques; une autre consiste à inverser la séquence et demander aux experts et aux décideurs de contribuer à la conception du modèle dès le départ.

7.6 **IRCC devrait régulièrement examiner et ajuster ses systèmes automatisés** pour s'assurer qu'ils reflètent la réalité sur le terrain et fonctionnent conformément aux priorités stratégiques du jour. Des boucles de rétroactions sont essentielles à une conception responsable et au contrôle de la qualité. Par exemple, la rétroaction des agents de première ligne doit se rendre aux concepteurs de systèmes pour qu'IRCC puisse continuellement réévaluer si les conclusions d'un système sont fiables. **Il est essentiel que des agents continuent d'examiner une portion des demandes manuellement.** Ainsi, les capacités des agents ne s'atrophieront pas et les connaissances sur les conditions d'un pays et les tendances en matière de fraude resteront à jour

7.7 **Les experts techniques (et leurs gestionnaires) sont en position de confiance et ont la responsabilité particulière d'évaluer en profondeur la logique d'affaires et les enjeux potentiels qui peuvent se manifester en raison de l'automatisation.** Il est essentiel que le Ministère reçoive des évaluations objectives et des conseils francs. Avec la montée des systèmes d'apprentissage machine, il faut faire extrêmement attention de cerner, d'expliquer et d'atténuer les risques potentiels, surtout si des risques ne sont apparents que pour les personnes comprenant les aspects techniques de ces systèmes. Les gestionnaires doivent créer des conditions dans lesquelles les experts peuvent consulter leurs pairs, réévaluer les risques à mesure qu'évoluent les systèmes et signaler les dangers sans peur de représailles.

7.8 IRCC devrait déployer tous les efforts pour **donner à ses employés les compétences, les outils et le soutien dont ils ont besoin** pour innover. Par soutien, on entend le temps accordé aux employés pour qu'ils apprennent et l'accès à des bacs à sable ou des environnements de test espace pour qu'ils expérimentent de façon responsable.

Voir la Directive du SCT, parties 6.3.9-6.3.10 (Assurer une intervention humaine) et 6.3.5 (Formation des employés).

8. En plus de respecter le cadre de protection de la vie privée du Canada, IRCC devrait continuellement adopter des pratiques exemplaires en matière de protection de la vie privée dans un domaine qui évolue rapidement.

8.1 Les scientifiques de données, les concepteurs de programmes et les spécialistes en TI devraient parfaitement comprendre les enjeux touchant la protection de la vie privée, ainsi que les **droits et responsabilités associés à la collecte, l'utilisation, la divulgation et la rétention de renseignements personnels** (lien interne), ainsi que **les normes et pratiques exemplaires connexes.** IRCC devrait

envisager d'élaborer des outils et des procédures pour aider le personnel à évaluer et répondre aux besoins en matière de protection de la vie privée dans cette ère de données massives.

8.2 Le Ministère devrait adopter une **approche proactive afin de se préparer à d'éventuelles infractions à la vie privée ou enquêtes visant sa gestion des renseignements personnels** dans le cadre d'un système algorithmique. Cela comprend la conformité avec des lignes directrices relatives aux infractions à la vie privée et à la sécurité, et des mesures pour veiller à ce que la formation soit à jour. Dans le cas d'une plainte ou d'une vérification du Commissariat à la protection de la vie privée, IRCC devra être en mesure de présenter des copies de tout document décrivant les démarches faites pour se conformer aux exigences en matière de protection de la vie privée, voire des démarches pour aller au-delà de la conformité, s'il y a lieu.

Voir la Directive du SCT, partie 6.3.3 (Qualité des données).

9. IRCC doit soumettre les systèmes à une surveillance continue pour s'assurer qu'ils sont techniquement adéquats, conformes aux exigences juridiques et en matière de politiques, justes et fonctionnels comme prévu.

9.1 Pour assurer la légitimité des systèmes, il est important d'encourager le **contrôle externe** de leur conception. Lorsque possible, les nouveaux modèles algorithmiques devraient être accessibles aux experts externes, comme un comité consultatif, à des fins de vérification, d'essai et d'examen.

9.2 Des processus de contrôle de la qualité devraient être établis dès de début et effectués régulièrement. Si les essais sont menés à l'interne, il faudrait documenter les méthodes et les hypothèses mises de l'avant, ainsi que les résultats qui en découlent, puis les rendre accessibles au comité consultatif. Soumettre les systèmes décisionnels automatisés à plus de mécanismes de contrôle avant qu'ils soient mis en œuvre, et sur une base continue par la suite, tout en assurant une approche plus légère pendant la phase d'exploration permettra d'en faire une surveillance adéquate sans freiner l'innovation.

Voir la Directive du SCT, parties 6.3.2 (Évaluer et surveiller les résultats) et 6.3.4 (Examen par les pairs).

10. IRCC doit toujours expliquer clairement ses décisions relatives aux demandes de clients.

10.1 Peu importe le degré auquel IRCC s'appuie sur ses modèles algorithmiques pour soutenir sa prise de décisions, les demandeurs auront toujours le droit de comprendre le fondement des décisions prises quant à leur demande. Ces explications sont essentielles pour permettre au demandeur de profiter des options de recours et pour assurer la responsabilisation du gouvernement et la confiance du public en général. Il peut être difficile sur le plan technique d'expliquer le fonctionnement des systèmes algorithmiques, mais **le public ne devrait pas avoir besoin de comprendre ce dernier pour qu'on lui explique les décisions automatisées**. Si une explication procure peu de renseignements utiles, alors elle offre peu de valeur.

10.2 Une des façons de rendre des explications compréhensibles est de limiter le rendement d'un système aux fonctions qui sont facilement explicables. Toutefois, cela peut signifier qu'il faudra

gaspiller le potentiel des systèmes d'IA qui peuvent comprendre les données complexes de manières dont les humaines ne peuvent pas. **Au lieu d'imposer des limites strictes sur ce que les outils peuvent *faire*, IRCC devrait tenter de trouver de nouvelles et meilleures façons d'*expliquer* ce qu'ils peuvent faire.**

10.3 Étant donné l'interdiction des systèmes « boîtes noires » pour la prise de décisions administratives complètes, l'utilisation des systèmes de soutien automatisé à la prise de décision devrait rendre les décisions **plus facilement explicables, pas moins**. Par exemple, on peut utiliser des pistes de vérification tenant compte de tous les facteurs et étapes dans le processus de prise de décisions, au besoin, pour générer et consigner systématiquement les raisons d'une approbation ou d'un refus.

10.4 Une **piste de vérification** est vitale pour respecter le droit d'un demandeur de comprendre le fondement de la décision prise concernant sa demande et peut être utile pour satisfaire un tribunal dans le cas d'une contestation juridique. Tout nouvel outil incapable de **générer automatiquement une piste de vérification sur le cheminement de la prise de décisions administratives** n'est pas prêt à la mise en œuvre.

Voir la Directive du SCT, partie 6.2.3 (Parvenir des explications après les décisions).

11. IRCC doit être transparent quant à son utilisation de l'IA. Il doit fournir un accès valable au système tout en protégeant la sécurité des Canadiens.

11.1 **IRCC doit trouver des façons de faire rapport, de façon significative, sur le fonctionnement des systèmes dans leur intégralité**. Il faut au minimum :

— Que les clients comprennent comment leur demande sera traitée et leurs renseignements personnels utilisés;

— Que les Canadiens comprennent comment IRCC dépense l'argent des contribuables pour faciliter l'entrée de familles, d'amis et de visiteurs légitimes tout en protégeant le public contre les menaces à la santé et à la sécurité;

— Que les intervenants, les médias et les experts techniques aient une base pour faire une critique éclairée de l'approche du Canada. Cela signifie qu'IRCC devrait partager de manière proactive des renseignements sur le fonctionnement de ses modèles et, plus important encore, les façons dont le Ministère utilise – et n'utilise pas – les nouvelles technologies pour soutenir la prise de décisions.

11.2 Il est dans l'intérêt du Ministère de ne pas divulguer trop d'information sur ses systèmes de décisions pour des raisons d'intégrité des programmes. En divulguant trop d'information, il pourrait faciliter la tâche aux personnes voulant manipuler la prise de décision ou « contourner le système ». Il doit fournir des explications sur le fonctionnement des systèmes et sur les décisions qu'ils aident le Ministère **suffisamment détaillées pour indiquer de manière générale à un client comment une décision est prise, mais pas trop détaillées au point de compromettre l'intégrité des programmes et, par extension, la sécurité des Canadiens.**

11.3 **IRCC devrait viser à concevoir des systèmes, dans la mesure du possible, en s'appuyant sur des points de données et des facteurs qui sont difficiles à falsifier ou à inventer**. Par exemple, des dossiers bancaires sont plus difficiles à falsifier qu'un motif de voyage, que le demandeur n'a qu'à déclarer. Nous devrions nous appuyer sur les faits objectifs, le plus possible, pour atténuer tout risque de fraude. Le degré d'importance accordé aux différents éléments de données dans les algorithmes d'IRCC devrait tenir compte du taux de tentative de fraude dont ils ont été la cible par le passé ou de la probabilité qu'ils soient ciblés à nouveau à l'avenir. Par exemple, on accorderait un degré d'importance élevé aux facteurs de risque à la sécurité.

Voir la Directive du SCT, parties 6.2.1-6.2.2 (Parvenir un avis avant les décisions) et 6.2.6 (Publication du code source).

12. Le fait qu'IRCC utilise des systèmes automatisés ne doit pas réduire la capacité d'une personne à miser sur des recours.

12.1 Lorsqu'un client conteste une décision qui a été prise avec le soutien d'un système automatisé, et qu'IRCC a le pouvoir et accepte d'examiner la décision, **alors ce sera un humain qui effectuera cet examen.** De même, en vertu de la *Loi sur la protection des renseignements personnels*, les individus ont le droit de recours – par l'intermédiaire du Commissariat à la protection de la vie privée – s'ils sont de l'avis que leurs renseignements personnels ont été traités de façon inappropriée.

12.2 Toutefois, **IRCC ne devrait pas offrir proactivement aux clients, au début du processus de demande, l'option qu'un agent examine leur dossier et prenne une décision à cet égard au lieu du système automatisé**. L'objectif d'IRCC est de mettre en œuvre des systèmes responsables, efficaces et efficients. En permettant aux clients de choisir d'éviter les systèmes, nous compromettrions la capacité du Ministère d'assurer un traitement rapide et cohérent, et cela laisserait entendre que les systèmes automatisés offrent un service inférieur.

Voir la Directive du SCT, partie 6.4 (Recours).

PARTIE II – LE MANUEL DES RESPONSABLES DE L'AUTOMATISATION

Ce guide pratique est conçu pour aider le personnel d'IRCC à répondre aux questions clés qui doivent être considérées aux différentes étapes des projets d'IA et d'automatisation – des premières phases de l'exploration à la surveillance continue une fois que le système est fonctionnel.

Il est important de consigner les décisions clés concernant l'utilisation de l'IA et de l'automatisation afin de démontrer que vous avez soigneusement pris en compte une série de facteurs et cerné en cours de route des stratégies d'atténuation adéquates. Il peut être utile de prendre note des observations, des analyses et des liens vers des documents clés tout au long des phases de conception, de développement et de déploiement afin d'étayer les réponses que vous devrez fournir lorsque vous effectuerez une évaluation de l'incidence algorithmique (EIA) et aussi en cas de litige. Vous trouverez de plus amples renseignements concernant l'outil d'EIA et les exigences dans la section B, *Concevoir votre système.*

Qui doit lire cette section?

- *Les personnes et les équipes qui envisagent de recourir à l'IA et à l'automatisation pour appuyer la prise de décisions administratives. Il peut s'agir de scientifiques des données, de responsables des programmes et des politiques, de professionnels en matière de gouvernance des données et de protection de la vie privée, de spécialistes des TI et d'experts en orientation opérationnelle – le plus vraisemblablement à l'échelon des analystes, des agents, des conseillers et des gestionnaire.*

A. Explorer le soutien automatisé à la prise de décision à titre de solution

Cette section vous aidera à passer en revue les questions clés qui doivent être prises en considération dès le début d'un projet. Les réponses à ces questions devraient vous aider à déterminer si un système de décisions automatisé est une solution qui convient ou non au problème que vous tentez de résoudre, si vous disposez ou pas de ce dont vous avez besoin pour réussir, et si les avantages surpassent les risques.

Pertinence générale

Avez-vous clairement défini le problème en cause?

Le fait de consacrer du temps au problème avant de se précipiter sur les solutions possibles vous permettra d'éviter un « théâtre technologique » et de vous assurer que la solution que vous finirez par choisir répondra à un besoin clair. Cela peut également vous faire gagner du temps et vous aider à cerner les indicateurs pertinents pour mesurer le succès.

Tous les intervenants comprennent-ils le problème de la même façon? S'agit-il d'une question qui peut être résolue par l'IA et l'automatisation?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous épuisé toutes les options simples avant d'envisager le domaine complexe et litigieux du soutien automatisé à la prise de décision?
Dans certains cas, le recours à l'automatisation de tâches simples et répétitives (p. ex. en utilisant l'automatisation robotique des processus) ou de nouvelles approches quant à la répartition de la charge de travail pourrait permettre d'atteindre des gains d'efficacité similaires. Envisagez d'emblée plusieurs options.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

La nature de la décision ou l'étape précise dans le processus décisionnel est-elle propice à la prise de décision automatisée?

Au moment de déterminer si une tâche ou une étape dans le processus décisionnel doit passer de l'humain à la machine, vous devriez vous poser la question suivante : « dans cette situation, des esprits raisonnables pourraient-ils aboutir à une conclusion différente? » Si oui, alors l'automatisation n'est pas recommandée. À l'inverse, si l'analyse de décisions historiques montre que pratiquement tous les agents tireraient la même conclusion selon les faits dont ils disposent, l'automatisation devrait être poursuivie.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Votre projet requiert-il une attention particulière? Par exemple :

- Il s'inscrit dans un domaine sujet à l'examen minutieux du public (en raison des préoccupations touchant la vie privée, par exemple) ou à des litiges fréquents
- Les clients dans ce secteur d'activité sont particulièrement vulnérables
- Les enjeux des décisions sont très élevés; par exemple, vous envisagez des refus automatisés ou bien les résultats auront une incidence importante sur la vie des gens
- Cela aura une incidence importante sur les employés d'IRCC, que ce soit sur le plan de leur nombre ou de leur rôle
- Vous envisagez recueillir ou utiliser des données non traditionnelles ou potentiellement sensibles, comme des données associées aux médias sociaux des clients

Si vous avez répondu par l'affirmative à une ou plusieurs questions, votre projet nécessitera des consultations et une surveillance accrues.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Est-il probable que la technologie que vous envisagez utiliser soulève des inquiétudes du public? Par exemple :

- Cette technologie fait l'objet d'une couverture médiatique défavorable ou est associée à des résultats moins fiables dans le cas de certains groupes minoritaires (p. ex. reconnaissance faciale)
- Vous envisager utiliser un algorithme « boîte noire »

- *Comment savoir si vous avez une « boîte noire »? Le processus d'analyse ou de justification du système est inconnu ou son comportement est difficile à interpréter ou à expliquer.*

Si c'est le cas, il faudrait réfléchir à ce que vous pourrez faire afin que l'adoption de ces outils ne n'ait pas une incidence négative sur la confiance dans la capacité d'IRCC à prendre des décisions qui sont justes et fiables.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Diagnostics et évaluations d'impacts préliminaires

Quelles incidences, négatives et positives, l'automatisation proposée aura-t-elle sur les clients?

Par exemple, il est prévu que ce projet :
- Permette un traitement accéléré
- Exige que des clients qu'ils confient à IRCC davantage de renseignements personnels ou touche autrement leur vie privée
- Améliore ou dégrade la qualité des renseignements fournis aux clients quant aux décisions prises
- Fasse en sorte que certaines demandes font l'objet d'un examen moindre ou d'aucun examen par un décideur humain

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Votre système décisionnel automatisé aura-t-il un effet sur des groupes diversifiés?

L'analyse comparative entre les sexes (ACS+) est un outil analytique exhaustif qui permet d'évaluer pleinement les effets potentiels associés aux nouvelles initiatives et politiques gouvernementales sur divers groupes de personnes. Le « plus » prend acte de multiples facteurs qui se recoupent comme : le sexe et le genre, le statut d'autochtone, le revenu, la situation professionnelle, la situation familiale, le niveau de scolarité, le statut d'immigrant et le pays de naissance, l'âge, le statut de minorité visible, la langue, la région de résidence et l'état de personne handicapée.

L'application au tout début de votre projet de l'optique de l'ACS+ vous permettra d'évaluer les risques qui peuvent exposer des groupes divers à des effets préjudiciables. À ce stade, vous devriez vous concentrer sur la collecte de données désagrégées au sujet des clients, sur l'analyse de ces données pour en déterminer la qualité et en dégager les préjugés historiques et sur la vérification de vos hypothèses. L'IA peut également s'avérer utile pour établir les tendances pertinentes d'intersectionnalité que les humains pourraient ne pas remarquer. L'ACS+ entrera également en jeu à des stades ultérieurs de votre projet!

Il vous faut contacter dès que possible l'Unité de l'ACS+ d'IRCC, qui relève de la Direction générale des politiques et de la planification stratégiques, pour obtenir des conseils et suivre les formations offertes.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Quelles incidences, négatives et positives, l'automatisation proposée aura-t-elle sur le Ministère?

Par exemple :

- Réduire l'arriéré des cas existant
- Faire des économiques (à court, à moyen ou à long terme)
- Améliorer la cohérence ou la qualité générale des décisions
- Améliorer l'intégrité des programmes
- Exécuter des tâches que les humains ne seraient pas en mesure d'accomplir dans un délai raisonnable, permettant ainsi aux agents de se concentrer sur des tâches plus complexes nécessitant le jugement humain
- Augmenter le risque de litige ou d'infraction à la vie privée
- Augmenter (ou réduire) la validité des données recueillies et utilisées
- Changer le rôle des agents et des décideurs, y compris les compétences essentielles, les descriptions d'emploi ou les classifications des décideurs, ou encore changer la taille et la composition du personnel requis pour exécuter un programme

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous réfléchi à la réaction des membres du personnel de première ligne?

Il serait utile de les consulter et de consulter les ressources humaines. Même les propositions qui ne prévoient pas de grands changements au rôle des agents peuvent provoquer chez eux de l'incertitude et du stress. En veillant dès le début à communiquer les objectifs du projet, vous rassurerez le personnel, gagnerez du temps et des efforts plus tard.

Réfléchissez à la manière dont vous obtiendrez la rétroaction du personnel de première ligne à l'égard d'un outil d'automatisation proposé. Cela vous donnerait l'occasion de communiquer les objectifs d'un outil d'automatisation proposé ou permettrait de répondre aux préoccupations le plus tôt possible. La section sur l'adoption d'une approche axée sur les utilisateurs propose une orientation plus détaillée.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous cerné les risques et les répercussions juridiques potentiels?

Dès les premières étapes de votre projet, vous devrez établir un diagnostic des enjeux juridiques et signaler les répercussions potentielles à ce chapitre en vue d'un examen plus approfondi. Cette démarche doit porter sur tous les domaines du droit public (c.-à-d. droit administratif, droits de la personne et protection de la vie privée). Pour de plus amples renseignements sur la manière d'évaluer et d'atténuer les risques juridiques tout au long du cycle de vie de l'IA, consultez ci-après les *Considérations juridiques et conseils pratiques*.

L'Unité des services juridiques d'IRCC comporte une équipe d'avocats spécialistes des questions liées à l'utilisation de technologies comme l'IA et l'automatisation. Vous devez contacter les Services juridiques au début du processus et régulièrement par la suite pour obtenir des conseils relativement à votre projet.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Aurez-vous besoin de nouveaux pouvoirs politiques, juridiques ou réglementaires?

Le Ministère possède un vaste pouvoir d'utiliser des outils électroniques pour réaliser son mandat en matière de prise de décision liée à l'immigration et aux passeports (voir l'article 4.1 de la *Loi sur l'immigration et la protection des réfugiés* [LIPR] et les modifications à l'article 2.2 du *Décret sur les passeports canadiens*), mais vous pourriez avoir besoin de nouveaux pouvoirs si votre projet entraînerait l'automatisation d'autres secteurs d'activités ou des changements importants à ce qui suit : exigences de programme (par exemple les critères de recevabilité) ; résultats de programme (par exemple, quels clients sont approuvés) ; collecte, utilisation ou communication de renseignements personnels; rôle des agents dans le processus de prise de décisions. Même si vous avez les pouvoirs juridiques pour mettre en œuvre l'automatisation proposée, des changements réglementaires pourraient s'avérer nécessaire pour établir les précautions ou pour préciser exactement comment et quand l'automatisation sera effectuée.

Il se peut que vous deviez travailler avec les affaires juridiques et les responsables des politiques pour obtenir de nouveaux pouvoirs ou pour mettre en œuvre des changements réglementaires au moyen d'un mémoire au Cabinet ou d'autres instruments.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

IRCC possède-t-il des données d'une portée et d'une qualité suffisantes pour que votre proposition soit viable?

L'établissement d'un diagnostic touchant la gestion des données vous permettra non seulement de cerner et d'atténuer les risques juridiques, mais vous aidera également à évaluer la faisabilité de votre projet. Ce faisant, vous devriez considérer la gestion des données en ce qui touche la collecte, l'utilisation, la communication, la conservation et la sécurité des données, y compris des renseignements personnels (qui pourraient nécessiter une évaluation des facteurs relatifs à la vie privée). Vous devriez également vous demander si vous devez élaborer une stratégie en matière de données aux fins du projet.

Voici quelques-unes des questions que vous devriez prendre en compte :

- Y a-t-il suffisamment de données structurées de grande qualité pour appuyer des conclusions fiables? Quelles sont les limites des données? Sont-elles exactes, à jour et véritablement représentatives, et contiennent-elles des préjugés historiques qui pourraient être maintenus? Vous devrez faire évaluer la qualité de vos ensembles de données avec l'aide du leadership de la fonction de gouvernance des données ministérielle et de la collaboration entre les opérations et les TI. Cette évaluation devrait être approuvée par un groupe de gestion de la gouvernance des données, comme le comité de direction des données.
- Combien de vos points de données sont des intermédiaires pour les critères ou résultats réels que vous visez? Tous? Vous aurez besoin d'indicateurs éprouvés indiquant que les demandeurs sont conformes aux objectifs ou exigences du programme.

— Pouvez-vous utiliser les données? Le modèle fait-il des inférences quant à des renseignements personnels qui ne sont pas fournis par le demandeur? Sont-elles invasives ou posent-elles des risques élevés au droit à la protection de la vie privée des personnes? Selon quel fondement utiliseriez-vous les données dans votre analyse dans une application du monde réel? Avez-vous besoin d'un accès à de l'information qu'IRCC recueille déjà (il faudrait examiner l'utilisation interprogrammes de l'information déjà recueillie pour un programme), ou cherchez-vous à tirer des données externes, comme celle que détiennent les partenaires fédéraux, provinciaux ou internationaux?
— De quelle manière les données seront-elles recueillies et stockées, par rapport à d'autres systèmes et biens de données d'IRCC?

Vous devrez obtenir des conseils auprès de la Direction générale du dirigeant principal des données, de la division d'AIPRP et possiblement des Services juridiques pour déterminer si la collecte, l'utilisation ou la divulgation des données que vous proposez sont conformes aux cadres de protection de la vie privée régissant le Ministère ou si vous devrez obtenir de nouveaux pouvoirs. L'Unité de l'ACS+ peut également fournir des conseils concernant la collecte et l'affichage de renseignements liés au sexe ou au genre.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous évalué les besoins en matière de protection de la vie privée liés à votre projet?

De façon générale, vous devrez examiner quels renseignements personnels peuvent être utilisés dans le cadre de votre projet et comment cette utilisation peut s'effectuer en mettant en jeu le moins d'intérêts liés à la vie privée. Vous devrez également évaluer les flux de données afin de vous assurer que la collecte, l'utilisation et la communication de renseignements personnels satisfont aux exigences juridiques et politiques applicables. En vous familiarisant avec les principes du respect de la vie privée dès la conception, vous pourrez réfléchir aux considérations et pratiques exemplaires éventuelles en matière de respect de la vie privée qui concourront à la solidité de votre proposition.

La Direction générale des affaires coopératives intégrées d'IRCC a élaboré de nouveaux outils et procédures internes pour permettre à son personnel d'adopter à l'égard de ses projets une approche qui tient compte de la vie privée, notamment les projets faisant intervenir l'IA, l'analytique avancée et l'automatisation.

La réalisation d'une **Analyse des besoins en matière de protection de la vie privée** (lien interne) est la première étape à suivre afin de déterminer les mesures que vous devrez prendre pour atténuer et gérer adéquatement les risques d'atteinte à la vie privée associés à votre projet. Les questions mettent l'accent sur les autorisations légales pour la collecte de données, les types précis de renseignements personnels qui seront recueillis ou créés, l'identité de ceux qui stockeront et consulteront ces renseignements personnels et les modalités d'utilisation des données (p. ex. informatique en nuage, robots conversationnels, analytique, IA, moissonnage du Web, couplage de données). Une fois que vous aurez rempli le formulaire, un conseiller en AIPRP l'examinera et pourrait recommander :

– l'élaboration d'une nouvelle Évaluation des facteurs relatifs à la vie privée ou EFVP (lien interne);
– la modification ou la mise à jour d'une ou plusieurs EFVP existantes;
– d'autres travaux relatifs à la protection de la vie privée, selon les besoins.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Formation

Avez-vous tiré profit de la formation offerte sur les technologies fondées sur les données et sur leurs effets?

Cela peut aller de cours sur le gouvernement numérique – comme ceux offerts par l'Académie du numérique de l'École de la fonction publique du Canada – à des formations sur la protection de la vie privée et le savoir-faire en matière de données. Les formations sur l'ACS+ vous permettront également d'examiner les questions entourant les préjugés et de cerner les mesures à prendre pour éviter les conséquences négatives imprévues sur divers groupes.

Même si tous les membres d'une équipe travaillant sur un système de décisions automatisé n'auront pas à acquérir le même niveau d'expertise dans tous les domaines, la *Directive sur la prise de décision automatisée* du SCT exige qu'une formation appropriée concernant la conception, la fonction et la mise en œuvre soit fournie aux employés afin qu'ils soient en mesure d'examiner, d'expliquer et de surveiller le fonctionnement d'un tel système.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Mobilisation des partenaires et des intervenants

Avez-vous réfléchi à propos des experts internes qui devront être mobilisés dans le cadre du projet?

Une vaste expertise est requise dans de nombreux domaines pour garantir l'utilisation responsable de l'automatisation dans la prise de décisions. Il vous sera utile de réunir un groupe multidisciplinaire pour aborder simultanément les enjeux politiques, juridiques, informatiques, opérationnels et les questions relatives aux données. Cela vous aidera également à cerner les stratégies pertinentes d'atténuation des risques (p. ex. les préjugés dans votre système automatisé), à mieux sensibiliser l'ensemble du ministère à votre projet et à faciliter l'harmonisation avec d'autres initiatives connexes.

Vous devrez probablement faire appel à des experts issus des équipes suivantes :

- **Centre de solutions en analytique avancée (CSA²), Direction générale de la planification et du rendement des opérations** – dans la plupart des cas, l'équipe du CSA², composée d'experts internes en sciences des données appliquées, sera chargée de concevoir et de mettre à l'essai des modèles de soutien automatisé à la prise de décision.
- **Services juridiques** – les avocats spécialisés dans les technologies peuvent fournir des conseils sur les exigences des programmes, la protection de la vie privée, la Charte, le droit administratif, la propriété intellectuelle et les risques de litiges. Aux termes de la Directive du SCT, vous devez consulter les services juridiques dès la conception du projet afin de vous assurer que l'utilisation du système de décisions automatisé est conforme aux exigences juridiques applicables.
- **Division de l'AIPRP, Direction générale des affaires corporatives intégrées** – l'Unité de la protection des renseignements personnels, de la politique et de la gouvernance peut fournir des conseils sur les considérations et les exigences en matière de respect de la vie privée, et vous

indiquer notamment comment remplir un formulaire Évaluation des besoins en matière de protection de la vie privée, préciser si vous devez effectuer ou réviser une évaluation des facteurs relatifs à la vie privée (EFVP), un énoncé de confidentialité ou un fichier de renseignements personnels (FRP) existants. L'équipe peut également collaborer avec vous pour intégrer la protection de la vie privée à des projets individuels.

- **Équipe des politiques numériques, Politiques stratégiques et planification** – cette équipe peut donner des conseils sur les pouvoirs juridiques et politiques, l'orientation d'IRCC et du SCT portant sur l'utilisation responsable de l'IA et de l'automatisation, les liens avec d'autres initiatives du Ministère, etc.
- **Équipe de l'ACS+, Direction générale des politiques et de la planification stratégiques**
- **Une direction générale responsable des politiques relatives au programme** – selon le cas d'utilisation, il est possible que vous deviez travailler avec les responsables des domaines politiques de programme pour assurer l'harmonisation de votre initiative avec les futures orientations politiques, et rechercher de nouveaux pouvoirs (p. ex. Direction générale de l'immigration, Direction générale de la citoyenneté, Direction générale des politiques de l'établissement et de l'intégration).
- **Direction générale du dirigeant principal des données** – en tant que responsable ministériel de la « qualité des données par la conception », cette direction générale peut vous aider à évaluer la qualité des données et leur pertinence pour votre projet, et vous fournir des conseils en matière de gouvernance des données (p. ex. qualité, gestion et cycle de vie des données et glossaires opérationnels). Cette équipe devrait participer à la conception, au développement et à la mise en œuvre des systèmes de soutien automatisé à la prise de décision.
- **Orientation du programme d'immigration, Orientation sur les risques pour l'intégrité et réseaux de traitement** – ces équipes peuvent fournir des connaissances pratiques sur le processus de prise de décisions, le rôle des agents, les orientations connexes et les incidences opérationnelles liées à votre projet.
- **Direction générale de la stratégie numérique** – cette équipe peut vous aider à évaluer comment votre initiative s'inscrit dans la stratégie numérique en matière de TI et les capacités opérationnelles d'IRCC.
- **Communications** – cette équipe peut vous aider à déterminer si des communications au public sont nécessaires concernant votre projet.
- **Direction générale de l'expérience client** – cette équipe peut appuyer le développement d'une conception centrée sur la personne de bout en bout et des tests de convivialité liés à votre projet.
- **Direction générale des ressources humaines** – cette équipe peut évaluer si un nouveau système automatisé nécessitera des modifications aux descriptions de travail génériques, aux énoncés des critères de mérite ou aux classifications de poste. Elle peut aussi fournir des conseils sur quand et comment consulter les syndicats concernés.
- **Opérations de TI, Secteur de la transformation et des solutions numériques (STSN)** – cette Direction générale est responsable de la conception et de l'architecture des systèmes et fournit des conseils pour répondre aux besoins du ministère en matière de sécurité informatique ainsi que pour veiller au respect des lois, des politiques et des normes applicables. Vous devrez faire appel à l'Unité de la sécurité des TI durant la phase de conception pour entamer un processus d'évaluation et d'autorisation de sécurité, obligatoire au titre de la Politique sur la sécurité du gouvernement.

Vu l'ampleur et la complexité potentielles des propositions d'automatisation, vous gagneriez peut-être aussi à réfléchir à la manière dont vous mobiliserez votre équipe de direction et d'autres hauts

fonctionnaires tout au long du développement de votre initiative. Cela permettra de s'assurer que les cadres dirigeants disposent de tous les renseignements nécessaires pour prendre des décisions au bon moment tout au long du processus.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous réfléchi à savoir si vous avez besoin de la participation d'autres ministères (p. ex. ASFC) ou des gouvernements provinciaux et territoriaux? Même si la présente politique cible les activités d'IRCC, votre projet pourrait avoir une incidence sur ses partenaires (qui peuvent aussi avoir en place des politiques et approches en matière d'analytique, d'IA et d'automatisation).

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Ferez-vous appel à des intervenants externes?

Il est essentiel de faire appel à des intervenants externes pour ce qui concerne l'utilisation par IRCC de l'IA et de l'automatisation dans la prise de décisions, afin de renforcer la confiance du public, de veiller à ce qu'il n'y ait pas de décalage important entre la manière dont IRCC utilise ces technologies et ce que le public juge acceptable et, en fin de compte, de développer notre capacité à progresser dans ce domaine de manière responsable.

Il ne sera peut-être pas pratique de consulter des intervenants externes à propos du développement de chaque algorithme ou système automatisé qu'utilise IRCC, mais vous pouvez les consulter sur les incidences potentielles. Il est important de tenir compte des perspectives de ceux qui seront touchés par les aspects automatisés du processus de prise de décision, particulièrement lorsque de nouveaux algorithmes ou systèmes sont intégrés à un tout nouveau programme ou à un changement substantiel de l'approche de prestation.

Les intervenants traditionnels d'IRCC sont toujours importants, mais vous devrez tenir compte d'un plus grand éventail d'intervenants, comme ceux qui travaillent dans le domaine de l'éthique des technologies ou de l'IA.

Il est recommandé que vous cherchiez délibérément les perspectives d'un groupe diversifié d'intervenants et que vous consigniez leurs perspectives comme vous le feriez dans le cadre du développement d'un changement important à une politique ou une loi.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Planification de la conception

Avez-vous bien prévu à votre budget des ressources pour le développement de systèmes d'itération et d'expérimentation d'analytique de données dans votre mémoire au Cabinet ou présentation au Conseil du Trésor (si nécessaire)?

Il faut consacrer du temps et de l'argent pour révéler de nouvelles connaissances opérationnelles au moyen de l'analytique avancée. Le processus pour créer un nouveau système de soutien aux décisions automatisé peut comporter des essais et des erreurs et en fixant un budget à cet égard, vous aurez la flexibilité de faire les itérations requises avant de réaliser le produit final visé. Une fois construit, un système automatisé a besoin de ressources pour sa maintenance et son renouvellement périodique.

Durant les premières étapes du projet, vous devrez probablement présenter l'approche proposée au Comité de la réception des demandes (CRD). Ce processus centralisé de réception des propositions d'investissement en GI/TI (et de projets sans rapport avec les GI/TI), dirigé par la Direction générale des projets, évalue la valeur opérationnelle et la priorité des nouvelles idées dans une perspective ministérielle. Il permet également de s'assurer que les nouveaux investissements appuient l'orientation stratégique définie par les priorités et la feuille de route de la transformation d'IRCC.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous réfléchi à la façon dont les besoins en ressources évolueront au fil du temps?

Il est courant d'avoir besoin de plus de ressources humaines et financières dès le départ, au moment où un régime d'assurance de la qualité strict est vital pour veiller au fonctionnement prévu du système. Il est aussi important de prévoir les coûts permanents relatifs à la maintenance, au suivi et au renouvellement du système.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Votre système de décisions automatisé sera-t-il conçu à l'interne ou par un fournisseur externe?

Dans la plupart des cas, la conception et le développement des modèles seront probablement réalisés à l'interne par le CSA[2]. Si les projets sont confiés à des tiers, vous devriez consulter le CSA[2] pour obtenir des conseils et tirer profit de son expertise technique. Vous devriez également consulter les lignes directrices du gouvernement du Canada sur le recours à des tiers fournisseurs d'IA et d'automatisation (p. ex. Lignes directrices sur les services et le numérique)

Services publics et Approvisionnement Canada et le SCT ont établi une liste de fournisseurs préqualifiés intéressés à fournir des solutions d'IA au gouvernement du Canada. Vous devrez également examiner attentivement les exigences du SCT en matière d'accès et de propriété des codes sources.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Approche axée sur l'utilisateur

Si les modèles de responsabilisation et d'intervention humaines sont des concepts familiers dans le domaine de la prise de décision automatisée, l'idée d'une approche axée sur l'utilisateur de bout en bout ne l'est pas autant.

Et pourtant, le défaut de tenir compte des besoins et des expériences des utilisateurs peut augmenter les risques qui vont de la responsabilité juridique (découlant d'une confiance excessive dans le système ou d'une entrave au pouvoir discrétionnaire) à une diminution des rendements tirés des investissements (parce que les nouveaux systèmes ne sont pas bien intégrés dans les pratiques opérationnelles existantes), en passant par des décisions de moindre qualité (parce que les utilisateurs ne comprennent pas suffisamment le système et ses résultats).

Avez-vous pris des mesures pour comprendre les besoins des utilisateurs et leur environnement actuel?

Avant d'entreprendre un projet qui pourrait avoir des répercussions considérables sur le rôle des agents dans la prise de décisions, vous devez bien comprendre l'environnement général des opérations en cause.

- Établissez qui seront les utilisateurs principaux et adressez-vous à eux directement pour comprendre leurs besoins. Règle générale, les utilisateurs des systèmes automatisés d'IRCC seront les agents de nos réseaux opérationnels. Utilisez une combinaison d'appels ouverts à des observations et de demandes ciblées pour recueillir des avis divers.
- Cernez les pressions, les difficultés et les limites existantes ainsi que les facteurs susceptibles d'appuyer ou d'entraver les changements. Il pourrait notamment s'agir de contraintes liées au temps, aux ressources ou à l'infrastructure des TI.
- Anticipez les répercussions (perçues et réelles) de l'automatisation sur les utilisateurs tout en reconnaissant qu'il est peu probable qu'ils soient tous affectés de la même façon.
- Familiarisez-vous avec les outils et les processus déjà utilisés (ou prévus) tout au long du processus de prise de décisions ainsi que dans les bureaux ou les secteurs d'activités ciblés. Cela vous permettra de développer la compatibilité et d'harmoniser votre système d'une manière qui ait un sens pour les utilisateurs.
- Prenez en compte vos propres croyances et préjugés potentiels en ce qui touche l'environnement opérationnel actuel et la manière dont ces facteurs pourraient affecter votre interaction avec les utilisateurs. Cet aspect doit être inclus dans votre analyse de l'ACS+ plus générale.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

B. Concevoir votre système

La présente section répondra à des questions comme celles-ci : *Comment assurons-nous l'équité dans notre conception? Comment protégeons-nous l'impartialité des décideurs? Comment répondons-nous aux attentes des clients en matière de transparence et de protection des renseignements personnels?*

Vous devrez également fournir certaines des réponses lorsque vous réaliserez l'évaluation de l'incidence algorithmique (EIA) – des renseignements supplémentaires concernant l'EIA sont fournis ci-après.

Pertinence du modèle

Avez-vous mis à l'essai plus d'une option?

Il est peu probable qu'il n'y ait qu'une façon d'atteindre votre objectif. En explorant d'autres options, en établissant des modèles connexes et en les mettant à l'essai, vous pourriez confirmer votre hypothèse initiale et renforcer votre analyse de rentabilité ou découvrir une meilleure approche.

N'oubliez pas que le cadre de protection de la vie s'applique même pendant les étapes de la création du modèle de vos activités. La Division de l'AIPRP peut fournir des renseignements concernant des exigences précises et donner des conseils pour vous aider à adopter une approche tenant compte de la protection de la vie privée.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous élaboré un diagramme de processus détaillé qui permettra aux spécialistes des politiques, aux experts en protection de la vie privée, aux avocats et à d'autres personnes en dehors de votre domaine de comprendre le processus en question et comment, précisément, l'établissement de l'automatisation le modifiera?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Évaluation de l'incidence algorithmique (EIA)

Avez-vous effectué une EIA préliminaire?

L'EIA est un questionnaire conçu pour vous aider à évaluer et à atténuer les risques liés au déploiement d'un système de décisions automatisé. Il comprend environ 60 questions portant sur le projet, l'algorithme et les données ainsi que sur les mesures d'atténuation et de réduction des risques aux fins de la qualité des données, de l'équité procédurale et de la protection de la vie privée.

La réalisation d'une évaluation préliminaire au stade de la conception vous permettra d'anticiper les risques associés à votre projet, de cerner les mesures d'atténuation et de planifier le respect des exigences qui concernent les éléments suivants :

- Examen par les pairs
- Avis
- Intervention humaine dans les décisions

- Explications requises
- Mises à l'essai
- Surveillance
- Formation
- Planification d'urgence
- Approbation de la mise en marché du système

Pour déterminer les réponses du ministère aux questions de l'EIA, il faudra obtenir le point de vue de plusieurs experts internes. À cette fin, vous devriez faire appel à des représentants des politiques et des programmes, à des conseillers juridiques, à des spécialistes de la gouvernance des données et à des scientifiques des données et conserver les documents ou analyses à l'appui. Vous feriez bien aussi d'obtenir l'approbation de certaines directions générales (p. ex. approbation de la Direction générale du dirigeant principal des données des questions concernant les données d'entrée et la qualité des données).

Vous devrez mettre à jour l'EIA afin de tenir compte des changements importants apportés à votre modèle, et divulguer publiquement les résultats finaux.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Conception axée sur l'utilisateur

Avez-vous fait appel à des agents pour concevoir l'interface utilisateur de votre système et pour réfléchir à l'approche la plus efficace pour assurer l'intervention humaine?

Pour tirer le meilleur parti des outils d'automatisation, il est nécessaire de réfléchir à la fois à la manière la plus appropriée d'utiliser la technologie et à la meilleure façon d'assurer une intervention humaine. Une démarche axée en particulier sur l'interaction personne-machine permettra de renforcer la responsabilité, de protéger l'impartialité des décideurs et d'obtenir le bon équilibre entre le rôle des humains et le système. Le fait de faire appel directement aux utilisateurs pour la conception de ces processus (plutôt que de les en informer simplement après le fait) permettra de bâtir la confiance dans le système automatisé. Une co-conception efficace doit également s'appuyer sur une expérimentation rigoureuse et une approche itérative – c.-à-d. commencer à petite échelle, mettre à l'essai différentes approches, recueillir des observations, adapter la conception et recommencer.

- Essayez différentes manières de présenter les informations concernant la machine aux agents, en utilisant une approche rigoureuse pour mesurer l'impact des différentes conceptions. Par exemple, l'utilisation d'étiquettes à code couleur pour le triage des demandes peut susciter différents types de réponses comparativement à la présentation d'informations étiquetées avec un pourcentage ou un niveau de complexité. Des interactions personne-machine et des sorties machine mal conçues peuvent masquer les problèmes de biais d'induction liés à l'automatisation et entraver la prise de décisions.
- Veillez à ce que l'interface utilisateur et les sorties système encouragent les utilisateurs à exercer leur propre jugement.
- Collaborez avec les agents et d'autres membres de l'équipe multidisciplinaire pour définir les points de contact clés et les tâches nécessitant une intervention humaine. Vous devriez

également examiner les circonstances dans lesquelles un agent serait autorisé à passer outre une décision prise par le système.

— Veillez à concevoir et à dispenser une formation qui permette aux agents de comprendre pleinement les capacités et les limites du système.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Équité et non-discrimination

Quelles étapes suivrez-vous pour minimiser les préjugés involontaires?

Des algorithmes mal conçus peuvent introduire des préjugés, mais dans la plupart des cas, les préjugés proviennent des données utilisées pour l'analytique ou la formation des modèles d'apprentissage machine. Les concepteurs de programmes doivent être conscients des faiblesses dans les données au cœur de leur projet. Parmi les techniques visant à atténuer les préjugés (article en anglais) dans les données, notons la consultation d'experts pour veiller à ce que les données soient situées dans le bon contexte (par exemple, IRCC a-t-il changé certaines définitions opérationnelles au fil du temps?) et la formation contradictoire (en anglais seulement) exigée pour les modèles d'apprentissage machine qui pourrait corriger toute injustice perçue. Vous gagneriez aussi à faire appel à des partenaires ou à des experts externes pour qu'ils examinent les ensembles de données que vous proposez et vous aident à cerner les lacunes ou à relever des préoccupations éventuelles. La mise à l'essai rigoureuse de l'assurance de la qualité peut aussi aider à détecter les tendances qui suggèrent les préjugés et les corriger immédiatement.

Peu importe les détails d'un projet donné, il est toujours conseillé de réunir différentes perspectives et identités au sein de votre équipe de développement. Il existe de nombreux exemples de projets d'IA qui ont connu des ratés en raison du manque de diversité dans le groupe d'experts. Si vous ne tenez pas compte des circonstances uniques des différents groupes de gens concernés, cela risque de nuire aux populations que votre système est pourtant conçu pour aider. Pour vous protéger contre des préjugés involontaires, vous devez avoir un point de vue d'ACS+ pour votre analyse de données et les premiers travaux de conception. Vous devriez vous efforcer, le plus possible, d'avoir au sein de votre équipe des membres qui ont différentes vues du monde, peu importe le sexe, le genre, l'âge, l'ethnie, la langue, la ruralité, la capacité ou d'autres facteurs d'identité.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Comment votre système sera-t-il conçu pour assurer l'équité procédurale?

Afin de protéger ou d'améliorer l'équité procédurale, vous devrez réfléchir attentivement à la façon dont l'ajout d'un soutien automatisé à la prise de décision changera le traitement des demandes et la prise de décisions connexes. Par exemple :

– Si vous envisagez un modèle prévisionnel, vous devriez réfléchir à comment il informera les agents sans entraver l'exercice de leur pouvoir discrétionnaire.
– Veiller à ce que tous les facteurs pertinents sont pris en compte.

– Si votre système tirera de l'information ailleurs que celle fournie par les demandeurs, réfléchissez à la façon dont les clients en seront avisés. Les demandeurs ont le droit de savoir quelles autres preuves externes ont été utilisées lors de la prise de décisions, en partie pour qu'ils aient la possibilité de corriger toute information erronée.
– Déterminer les raisons qui seront fournies lorsqu'une décision est prise.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Explicabilité et transparence

Comment assurerez-vous que des explications significatives sont présentées aux clients?

La *Directive sur la prise de décision automatisée* du gouvernement du Canada exige que des explications significatives soient fournies aux personnes concernées sur la façon dont une décision a été prise et la raison pour laquelle elle a été prise. Les critères particuliers varient selon le niveau d'impact associé au système automatisé.

En règle générale, les explications sur les décisions automatisées devraient : (1) aider le client à comprendre pourquoi une décision donnée a été obtenue et (2) lui fournir les motifs pour contester la décision s'il le désire. Les concepteurs de programme devraient aussi déterminer s'il sera possible d'indiquer, de façon globale, ce que le client devrait changer pour qu'il obtienne le résultat qu'il souhaite à l'avenir. On appelle cette approche « explication contre-factuelle ».

IRCC n'a pas encore mise en œuvre les refus automatisés des demandes des clients. S'il en venait à le faire, les concepteurs de programme devraient porter attention à ce qui suit :

– Pour les décisions administratives prises par des systèmes automatisés, ou avec l'aide de ceux-ci, il faut au moins pouvoir les expliquer aussi aisément que les décisions traditionnelles prises seulement par des humains. C'est pourquoi les algorithmes « boîtes noires », comme des réseaux neuronaux avancés, ne sont pas recommandés comme moyen principal de déterminer si une demande est approuvée ou non.
– Les explications ne doivent pas compromettre l'intégrité du programme. Heureusement, il n'est généralement pas nécessaire d'être transparent au point de divulguer le fonctionnement interne d'un système d'IA. Plutôt, on doit fournir des explications qui répondent à des questions comme celles-ci : Quel est le fondement de la décision? Quels sont les principaux facteurs? En changeant un facteur particulier, aurait-on obtenu une décision différente? Pourquoi deux dossiers similaires obtiennent-ils des décisions différentes?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Étant donné le besoin de transparence, particulièrement dans le contexte des contestations juridiques, proposez-vous des règles opérationnelles qui sont justifiables et rendent la tâche difficile aux gens qui veulent les manipuler?

— Les modèles et les règles doivent aborder les exigences de programmes de sorte que les décisions soient fondées sur ces exigences et sur de l'information pertinente et fiable.

- Concevez votre système, dans la mesure du possible, en fonction de facteurs ou de points de données qui sont difficiles à falsifier ou à inventer. Il suffit aux clients d'indiquer la plupart des renseignements sur leur demande, sans qu'IRCC puisse l'authentifier. Toutefois, il est possible de vérifier d'autres points de données, comme les dossiers bancaires ou des visas ayant déjà été accordés. En privilégiant ces points de données, on contribue à l'atténuation du risque que le demandeur qui connaît le fonctionnement des systèmes d'IRCC mente pour contourner les règles.
- Dans la pondération des variables dans un algorithme, tenez compte du degré auquel chaque variable a pu être la cible de fraude par le passé ou sera potentiellement ciblée à l'avenir.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Vie privée

Comment votre système assurera-t-il le respect des droits de protection de la vie privée des clients et des Canadiens?

Tenez compte des principes de protection de la vie privée à toutes les étapes du développement pour intégrer des mesures appropriées dès le début.

Dans votre approche à l'utilisation des renseignements personnels à des fins légitimes, sans violation des droits des personnes ou des groupes, vous devriez :

- Comprendre la provenance des données
- Limiter la collecte de renseignements personnels à ce qui est nécessaire
- Assurer l'exactitude des données
- Réduire les données marginales, redondantes ou inutiles
- Prendre des précautions pour éviter que des données anonymisées (ou pseudoanonymisées) soient de nouveau personnalisées et prévenir la collecte de données non autorisée
- Protéger les données contre l'accès non autorisé et la divulgation accidentelle
- Mettre en place des plans pour atténuer l'incidence dans le cas d'infraction à la sécurité, d'accès non autorisé ou de divulgation accidentelle

Vous devriez élaborer des politiques de données et les procédures de protection de la vie privée transparentes afin que votre projet permette aux personnes ce qui suit :

- Comprendre les données sont recueillies (et en assurer la gestion), comment elles sont utilisées et divulguées et pendant combien de temps elles sont conservées
- Donner leur consentement significatif quant à la collecte de données automatique et pour tout partage de données après coup
- Examiner et obtenir leurs renseignements personnels et corriger toute erreur

Un expert en matière de protection de la vie privée travaille au sein des équipes d'analytique avancée pour garantir la prise en compte de la protection de la vie privée à chaque étape des projets d'analytique avancée. En plus d'appuyer un processus plus transparent, cela signifie que les experts d'IRCC en matière de protection de la vie privée sont mieux à même de fournir des conseils sur les défis et les opportunités uniques en ce qui touche l'IA et l'automatisation. La Direction de consultation du gouvernement du Commissariat à la protection de la vie privée du Canada peut vous fournir des conseils

sur les questions de vie privée, même si vous ne menez pas une évaluation des facteurs relatifs à la vie privée. Pour ce faire, communiquer avec la Division de l'AIPRP, le point de contact central pour traiter avec le Commissariat.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Travailler ouvertement

Vous êtes-vous demandé si des informations concernant les travaux auxquels vous vous consacrez devraient être communiquées à l'externe?

Même si l'utilisation de l'IA dans la prise de décisions relève d'un domaine émergent et en pleine évolution, il existe de fortes chances que d'autres équipes réfléchissent à des questions semblables et qu'elles soient confrontées à des problèmes semblables. Bien que les données, les renseignements personnels et les codes de nature sensible doivent encore être protégés, il est possible que d'autres types d'information liés à votre projet puissent être communiqués ouvertement et réutilisés par d'autres.

Si vous avez mené des recherches, rédigé des rapports ou élaboré des orientations sur l'IA et l'automatisation qui pourraient être diffusés sans compromettre la confidentialité, la sécurité ou l'intégrité des programmes, demandez-vous s'ils pourraient être mis à disposition sur le site Web d'IRCC ou sur une autre plateforme d'élaboration de source ouverte, comme GitHub ou le portail de données ouvertes du gouvernement.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Votre système utilisera-t-il un logiciel de source ouverte?

La *Politique sur les services et le numérique* exige que les ministères utilisent des logiciels et normes de sources ouvertes par défaut, sauf si l'option de source ouverte n'est pas accessible ou ne répond pas aux besoins de l'utilisateur.

S'ils n'utilisent pas des sources ouvertes, les ministères doivent privilégier les options commerciales non contraintes aux plateformes par rapport aux produits commerciaux exclusifs pour éviter la dépendance à une technologie particulière et pour que soient possibles la substitution et l'interopérabilité.

Si une application personnalisée est développée, par défaut, tout code source écrit par le gouvernement doit être publié en format ouvert sur les sites Web du gouvernement du Canada et les services désignés par le SCT.

Voir les exigences de la *Directive sur la prise de décision automatisée* portant sur l'accès aux composants et la publication du code source, ainsi que l'orientation présentée dans la Ligne directrice sur les services et le numérique.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Responsabilité

Qui sera le décideur officiel?

L'article 4.1 de la LIPR indique que les agents peuvent prendre des décisions ou faire des déterminations ou des examens au moyen de systèmes automatisés rendus accessibles par le ministre. Cela ne signifie pas que les systèmes eux-mêmes deviennent les décideurs officiels. Les humains conçoivent des architectures de système, des règles opérationnelles, des seuils de fiabilité et d'autres mécanismes similaires et ce sont donc eux qui sont ultimement responsables de chaque décision que prend le Ministère. Pour sa part, le ministre est responsable devant le Parlement de toutes les activités de son ministère.

Vous devrez déterminer qui approuvera les règles opérationnelles utilisées dans votre système de soutien aux décisions et si cette même personne sera le décideur officiel.

Si un système automatisé détermine certains éléments d'une décision administrative, mais pas tous, la personne qui rend la décision définitive sera-t-elle considérée comme le décideur officiel pour tout le processus décisionnel? Ou plutôt, certains sous-éléments de la décision seront-ils considérés comme des décisions en soi ayant leur propre décideur officiel (par exemple, la recevabilité par rapport à l'admissibilité)?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Comment votre système conservera-t-il une piste de vérification pour chaque décision?

Pour se préparer à des contestations juridiques potentielles, IRCC doit absolument établir une chaîne de preuves démontrant comment les systèmes dans leur ensemble fonctionnent, ainsi qu'une piste concernant chaque décision individuelle ayant été touchée par ce système.

- Le système automatisé a la capacité de générer automatiquement une piste de vérification exhaustive démontrant le cheminement de la prise de décision administrative ou vous conserverez un document électronique dans le dossier du demandeur
- Tous les points de décision clés peuvent être repérés dans la piste de vérification
- Tous les points de décision clés s'inscrivent dans la logique du système automatisé relativement aux lois, politiques ou procédures applicables
- Toutes les décisions sont consignées et accessibles par l'utilisateur du système, soit un évaluateur ou un vérificateur
- La piste de vérification générée par le système automatisé peut servir à générer une notification de la décision (y compris un énoncé des raisons ou une autre notification) au besoin
- La piste de vérification est-elle protégée contre la falsification (pour assurer la protection et l'intégrité des données)
- La piste de vérification indique précisément quelle version d'un système automatisé a été utilisée pour chaque décision qu'il appuie
- La piste de vérification inclut un historique exhaustif et imprimable des modifications, notamment : qui a créé le dossier (avec l'heure et la date); qui a modifié le dossier (avec l'heure et la date); qu'est-ce qui a été modifié; pour les questions de protection de la vie privée et les

questions commerciales confidentielles, qui a consulté le dossier (avec l'heure et la date); qui a pris la décision définitive (avec l'heure et la date)

– La piste de vérification commence par indiquer le pouvoir ou le pouvoir délégué prévu dans la loi
– La piste de vérification indique qui est le décideur autorisé
– La piste de vérification permet l'enregistrement de l'intervention humaine dans les processus automatisés, par exemple le nom de la personne autorisée à intervenir

Pour l'AVE et les prolongations du permis d'études et de la fiche de visiteur, l'onglet d'automatisation dans le Système mondial de gestion des cas (SMGC) permet aux utilisateurs de voir clairement toutes les sous-activités, y compris lorsque des demandées sont « approuvées » ou sont redirigées pour un examen manuel.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Rassemblez-vous des documents énonçant le développement et le déploiement du système dans son ensemble?

Cela devrait inclure les spécifications de la conception, les données de formation, les versions du système et toute autre information qui pourrait être utile pour respecter les exigences de la Directive du SCT ou dans le cas d'une vérification, d'une enquête, d'une action coercitive ou d'une procédure judiciaire.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Comment votre approche d'assurance de la qualité continue veille-t-elle à ce que votre système automatisé produise constamment les bons résultats?

Vous commencerez peut-être par veiller à ce que des humains examinent chaque décision passant par le modèle, puis réduirez-vous peu à peu la portion des décisions examinées par des humains. Vous devriez également vérifier s'il y a entrave involontaire de la capacité des agents, préjugés ou discrimination.

Au cours d'un exercice d'assurance de la qualité, vous devriez vous assurer de consigner par écrit votre processus et ses résultats, en donnant assez de détails pour permettre à d'autres de le faire ou de l'examiner en cas de vérification ou de contestation juridique à l'avenir.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Sécurité

Votre système est-il exposé à un niveau inacceptable de risque en matière de cybersécurité?

Il est essentiel aux fins de l'intégrité des systèmes, du respect des obligations en matière de confidentialité et de la confiance du public dans le système d'immigration de garantir la sécurité des environnements dans lesquels IRCC développe, met à l'essai et déploie des modèles IA.

Le cadre de gestion des risques liés à la sécurité des TI d'IRCC vise à préserver la confidentialité, l'intégrité et la disponibilité des renseignements et des biens de TI ministériels, à protéger l'authenticité des transactions électroniques et à assurer la continuité des services et des opérations. Ce cadre favorise les meilleures pratiques, promeut les principes de gestion continue des risques tout au long du processus de développement du cycle de vie des systèmes et jette des bases solides pour le processus d'évaluation et d'autorisation de sécurité (EAS) d'IRCC.

Le processus d'EAS (une exigence de la *Politique sur la sécurité du gouvernement*) fournit un niveau d'assurance fondé sur le degré de préjudice raisonnablement attendu en cas d'accès non autorisé ou de divulgation des données. En fin de compte, l'EAS sert de donnée de référence qui documente la diligence raisonnable déployée dans la mise en œuvre des contrôles de cybersécurité et appuie les principes fondamentaux de responsabilité, de respect de la vie privée, de continuité de la preuve, de transparence démontrable et de surveillance.

Ce processus doit notamment comprendre :

- Une évaluation des risques liée aux menaces pesant sur le système
- Un examen de la conformité aux politiques existantes en matière de cybersécurité et aux contrôles de sécurité définis
- La confiance dans la protection du système – y compris des données qu'il utilise – contre le piratage et la manipulation

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

C. Se préparer au lancement

La présente section répondra à des questions comme celles-ci : *IRCC est-il prêt à mettre en œuvre le nouveau système? Avez-vous mobilisé toutes les personnes concernées et le Ministère est-il prêt à répondre aux questions et à réagir aux interruptions imprévues?*

Le processus d'obtention de l'approbation finale en vue du lancement pourrait dépendre de l'ampleur des changements que vous proposez.

Transparence

Avez-vous terminé d'élaborer votre approche concernant la transparence envers le public?

IRCC devrait communiquer aux clients et aux Canadiens tous les cas (secteurs d'activité, fonctions particulières) où un système automatisé est utilisé pour soutenir la prise de décision. Des experts d'IRCC en ce qui concerne les questions juridiques et techniques, la protection de la vie privée et les communications devront être consultés pour déterminer quels renseignements devront être divulgués publiquement et les modalités de cette divulgation.

À propos de ces systèmes, vous devriez transmettre au public :

- Une description du système en langage simple, y compris des exemples de ce qui suit
  - la raison d'être du système
  - la justification de son utilisation
  - son rôle dans le processus de prise de décisions
  - les résultats possibles pour les gens
- Une description générale des données sur lesquelles reposent le système et leur provenance (y compris un lien vers des données de formation anonymes accessibles au public)
- Une description générale des types de données utilisées dans le modèle
- De l'information sur le rendement, comme les taux d'erreur ou d'exactitude, la répartition des clients touchés (p. ex. par pays ou par programme), les coûts ou la productivité par rapport à une option avec intervention humaine seulement

La communication intégrale du code source au public démontre une transparence élevée, mais la plupart des gens ne sauront pas ce qu'il signifie. Ainsi, ce qui peut être pertinent de révéler au public est différent de ce que l'on divulgue à une communauté d'experts techniques. Lorsque la loi le permet et il est faisable de le faire sans exposer les systèmes d'IRCC à de nouvelles vulnérabilités, IRCC devrait divulguer son code source à un groupe d'experts techniques, tout en fournissant au public une explication facile à comprendre sur le fonctionnement du système. Cette approche de « conception ouverte » a l'avantage de protéger le Ministère contre les cyberattaques ou les fraudes sophistiquées perpétrées par des entités malveillantes ayant des compétences techniques, équilibrant ainsi la transparence avec l'intégrité du programme. Le groupe d'experts qui examine le code source et la conception du système de façon générale produisent un rapport contenant une section destinée au public.

S'il existe des détails sur le système qui, selon vous, ne devraient pas être divulgués au public, vous devrez confirmer s'il est possible de protéger cette information. Par exemple, y a-t-il des exemptions convenables aux termes de la Loi sur l'accès à l'information ou d'autres mesures de protection

législatives, comme la protection pour la sécurité nationale (et, s'il y a lieu, aux termes de la Directive du SCT)?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous préparé un avis pour les clients les informant qu'une décision à l'égard de leur demande sera prise en totalité ou en partie par un système automatisé?

Il faut rédiger en langage simple des avis au point de collecte de renseignements personnels et les rendre facilement accessibles sur la page Web du programme. La *Directive sur les pratiques relatives à la protection de la vie privée* et la *Directive sur la prise de décisions automatisée du SCT fournissent des détails supplémentaires sur les exigences relatives aux avis de confidentialité.*

L'extrait suivant provient de l'avis de confidentialité qui figure dans le formulaire de demande de VRT :

> *Les renseignements personnels recueillis au sujet d'une demande et d'autres renseignements recueillis à l'appui d'une demande, peuvent servir à l'analytique informatique à l'appui du traitement des demandes et de la prise de décisions, incluant votre demande. Les renseignements personnels, y compris ceux provenant de l'analytique informatique, peuvent également être utilisés à des fins comme la recherche, les statistiques, l'évaluation des programmes et des politiques, la vérification interne, la conformité, la gestion des risques, l'élaboration de stratégies et la production de rapports.*

Pour obtenir des conseils sur les avis de confidentialité, veuillez contacter la Division de l'AIPRP.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous consulté votre conseiller en matière de communication?

- Avez-vous défini les mises à jour nécessaires de la page externe d'IRCC concernant la prise de décisions automatisée?
- Une stratégie de communication sera-t-elle nécessaire?
- Y aura-t-il une annonce?
- Avez-vous préparé des messages clés, des foires aux questions, etc.?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Responsabilité

Avez-vous établi une structure de gouvernance qui définit clairement les rôles et responsabilités des principaux partenaires qui exécuteront le projet, notamment au moment de passer de l'environnement d'essai à un programme permanent?

La *Directive sur la prise de décision automatisée* du SCT précise que le sous-ministre adjoint chargé du programme utilisant un système de décisions automatisé est responsable des exigences concernant les

évaluations des répercussions algorithmiques, les mesures de transparence, l'assurance de la qualité, les recours des clients et la production de rapports. Cependant, les rôles et responsabilités particuliers varieront en fonction des structures plus générales de gouvernance d'IRCC et des projets individuels.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous rempli toutes les exigences en matière de protection de la vie privée et examiné tous les risques relevés dans votre évaluation des facteurs relatifs à la vie privée ou par le Commissariat à la protection de la vie privée (par le biais de vos interactions avec la Division de l'AIPRP d'IRCC)?

Faites appel dès le début à la Division de l'AIPRP pour veiller à ce que des mesures d'atténuation des risques d'atteinte à la vie privée puissent être adoptées dès la conception de votre initiative, plutôt qu'à la fin.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous satisfait à toutes les exigences de l'évaluation et de l'autorisation de sécurité et avez-vous examiné tous les risques relevés dans votre plan d'action en matière de sécurité (PAS)? Avez-vous une lettre de la décision d'autorisation?

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous une stratégie pour que les gestionnaires de projet assurent le suivi des progrès, mesurent les résultats et en rendent compte, et appuient les vérifications ou évaluations ultérieures?

Assurez-vous d'avoir établi les indicateurs appropriés pour mesurer précisément l'impact de la ou des fonctions de soutien automatisé à la prise de décision de votre programme.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Examen par les pairs

Avez-vous consulté les bons experts capables d'examiner la conception de votre système?

Selon l'impact potentiel de votre projet, la *Directive sur la prise de décision automatisée* du SCT pourrait vous obliger à le faire. Même si elle n'est pas obligatoire, la consultation d'experts en IA ou en sciences des données (p. ex. du Conseil national de recherches ou de Statistique Canada) peut aboutir à des suggestions utiles pour améliorer le système.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

État de préparation des utilisateurs

Avez-vous fourni une formation et des instructions claires et conviviales aux agents?

La compréhension des forces et des limites du système permettra à ses utilisateurs de savoir comment l'utiliser de manière appropriée et en tirer le meilleur parti. Si les agents n'ont pas vraiment confiance dans le fonctionnement d'un système automatisé, ils risquent davantage de ne pas tenir compte des renseignements qu'il génère. D'un autre côté, si leur confiance est excessive ou qu'ils ne savent pas comment interpréter les résultats, les agents peuvent avoir tendance à se reposer sur le système, ce qui pourrait passer pour une entrave à leur pouvoir discrétionnaire.

- Élaborez du matériel de formation et demandez aux utilisateurs de vérifier la clarté et la pertinence du contenu.
- Rendez les documents de FAQ aisément accessibles aux agents et aux autres membres du personnel.
- De concert avec les bons experts, déterminez s'il est nécessaire de mettre à jour les instructions d'exécution du programme pour expliquer les rôles et les responsabilités des employés et la manière d'utiliser les renseignements provenant de ces systèmes.
- Établissez des boucles de rétroaction pour permettre aux agents de fournir des commentaires sur leur utilisation du ou des systèmes. Ce type de mécanisme permettra aux utilisateurs de partager leur expérience directe et de faire entendre leur voix en cas de difficulté.
- Élaborez, au besoin, des stratégies de gestion du changement. Il est important de ne pas sous-estimer l'impact que les systèmes automatisés peuvent avoir sur les individus, les équipes et l'ensemble de notre organisation. La Direction générale de la culture et de la gestion du changement peut vous aider à cerner les mesures appropriées pour favoriser une transition en douceur.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

État de préparation du système

Avez-vous dressé un plan d'urgence dans le cas d'une panne de système (à court terme ou à long terme)?

Voir les exigences du SCT relatives aux systèmes et processus d'urgence dans la *Directive sur la prise de décision automatisée.*

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

État de préparation des partenaires

Les partenaires touchés, internes et externes, vous ont-ils confirmé qu'ils sont prêts au lancement de votre système?

Demandez-vous si ces partenaires auront besoin de soutien, comme des documents d'information ou des produits de communications.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Recours

Les clients souhaitant contester une décision concernant leur demande ont-ils suffisamment de recours pour le faire? La mise en place de l'automatisation nécessite-t-elle des changements ou des parcours additionnels?

Lorsqu'un client conteste une décision qui a été prise avec de soutien de l'automatisation, et qu'IRCC est autorisé à l'examiner et accepte de le faire, le processus d'examen devrait faire en sorte que A) la personne examinant la décision se mette dans la peau du décideur initial, et B) la personne examinant la décision n'est pas contrainte par la décision initiale, ni n'est contrainte de tenir compte toute conclusion ou recommandation fournie par le système informatique qui a contribué à la prise de décision. Si le décideur initial et la personne examinant la décision utilisent la même technologie, il y a un risque qu'un défaut ou une erreur de la technologie se répercute d'un processus à l'autre.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

D. Une fois le système fonctionnel

La présente section répondra à des questions comme celles-ci : *Le système fonctionne-t-il comme prévu? Faut-il revoir l'un des facteurs concernés?*

Assurance de la qualité et surveillance

Procédez-vous continuellement à l'assurance de la qualité pour vérifier que votre système ne génère pas d'erreurs ou de résultats accidentels?

Par exemple, cela pourrait consister réserver à une portion suffisante de toutes les demandes à l'examen exclusif d'un agent (humain) ou à sélectionner au hasard des dossiers qui passent par le système automatisé aux décisions et à les mélanger à la charge de travail normale des décideurs humains. Dans les deux cas, vous devriez prendre des précautions pour que les agents ne sachent pas quelles demandes sont nouvelles et lesquelles sont des vérifications d'assurance de la qualité.

Vous devriez également surveiller le fonctionnement du système afin de détecter et d'éviter les risques en matière de droit administratif et de respect de la vie privée, et déterminer si des effets discriminatoires résiduels peuvent être décelés.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Vos résultats à ce jour permettent-ils de conclure qu'IRCC pourrait réduire sans risque le seuil de fiabilité du système pour réaliser des gains d'efficacité? Ou l'inverse?

Les intervenants internes, comme les Services juridiques et les secteurs des politiques relatives au programme pertinents, devraient prendre part à l'examen des résultats et aux modifications apportées aux règles du modèle.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Surveillez-vous la sécurité de l'environnement?

Cela devrait inclure une surveillance par rapport à de nouvelles menaces de cybersécurité et l'intégrité du système dans son ensemble.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Comparez-vous les résultats aux répercussions attendues que vous avez définies lors des étapes antérieures du projet – p. ex. lors de l'analyse ou du diagnostic lié à l'ACS+?

Cette approche vous permettra d'évaluer si les mesures d'atténuation des risques concernant diverses populations sont adéquates et peut vous aider à relever toute autre conséquence négative involontaire.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Avez-vous de nouvelles données qui pourraient améliorer la prochaine itération de votre système de soutien aux décisions?

Le cas échéant, examinez si de nouvelles données pourraient remplacer les intermédiaires avec des preuves de résultats plus directes.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Le modèle que vous avez établi a-t-il généré de nouvelles connaissances sur votre programme?

Par exemple, l'expression du système d'intérêt facilitée par le programme Entrée express a soulevé des occasions de moderniser le programme économique fédéral dans son ensemble. De plus, l'analytique de données sous-jacente au modèle prévisionnel de visa de résidence temporaire lié à la Chine a renforcé la justification du programme CAN+ (soit l'historique des voyages).

Vous devrez vous pencher sur des questions telles que celle de savoir si le nouveau système remplacera effectivement les exigences du programme établies dans la législation ou la réglementation par des règles administratives générées par les algorithmes.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Y a-t-il un cycle et un processus établis pour l'examen périodique des règles opérationnelles, de la qualité des données, des seuils de fiabilité et des gains de productivité?

Sans une approche établie à l'examen périodique, il y a un risque que le fonctionnement adéquat du système repose sur le maintien en poste de certaines personnes. IRCC doit constamment cerner les limites des algorithmes (et des données sous-jacentes) et en informer les décideurs, peu importe le taux de roulement du personnel ou le changement de rôle.

Les lettres d'autorisation de sécurité des TI doivent être renouvelées au moins une fois tous les cinq ans, ou lors de changements importants apportés au système.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Évaluez-vous la performance globale du système et présentez-vous des rapports sur les résultats?

Selon les objectifs que vous avez fixés au départ, vous feriez bien de considérer comment l'automatisation a permis de réduire les coûts, de gagner du temps, de détecter les risques pour l'intégrité du programme ou d'améliorer la satisfaction des clients.

Vous devrez également évaluer la manière dont le système contribue à l'atteinte des objectifs globaux du programme et en rendre compte (voir les exigences en matière de rapport énoncées dans la *Directive sur la prise de décision automatisée*).

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Commentaires des utilisateurs

Surveillez-vous l'utilisation des systèmes automatisés et recueillez-vous les observations des utilisateurs?

Comme les réactions humaines à un système automatisé peuvent varier dans le temps (et les humains eux-mêmes pouvant changer du fait de la rotation du personnel), les systèmes doivent faire l'objet d'une surveillance permanente pour éviter des résultats négatifs involontaires.

- Faites participer les utilisateurs aux processus d'assurance de la qualité continus. Une partie des demandes des clients devrait être traitée manuellement par les agents. Outre le contrôle de la cohérence entre les décisions des agents et celles des machines (p. ex. au moyen de tests à l'insu), les agents pourront maintenir les compétences nécessaires pour traiter des demandes de complexité variable.
- Tirez profit des connaissances des agents pour orienter les changements et les mises à jour du système. La prise en compte de l'intelligence humaine permettra au système de demeurer pertinent au regard de l'évolution des risques et des tendances propres à chaque pays. Avant la mise à l'échelle ou le passage d'une phase pilote à un état permanent, faites appel à un groupe d'utilisateurs représentatif afin de recueillir leurs réactions, leurs observations et les leçons qu'ils ont apprises.
- Surveillez les changements dans le comportement des utilisateurs et l'utilisation du système. Vous avez conçu le système en ayant l'utilisateur à l'esprit, avez procédé à des mises à l'essai et modifié les sorties machine et l'interface, et avez communiqué les renseignements pertinents aux agents. Cependant, il est possible que leur comportement change au fil du temps. À mesure que les agents s'habituent à interagir avec le système, ils peuvent avoir plus ou moins confiance dans les sorties machine. Pour détecter les biais d'induction involontaires liés à l'automatisation ou une confiance excessive dans les sorties système, envisagez de faire participer les utilisateurs à des groupes de discussion périodiques, d'analyser leur comportement en situation réelle et d'analyser les données à la recherche d'anomalies. Les utilisateurs devraient également être mis à partie dès que vous envisagez des changements qui pourraient avoir un impact sur la conception de l'interface ou des sorties système.
- Tenez les utilisateurs au courant en faisant circuler des rapports sur la performance du système. Cela permettra de démontrer une ouverture et une transparence et permettra aux utilisateurs de mieux apprécier les avantages (et les défis) offerts par le soutien automatisé à la prise de décision.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

Passage de la phase pilote à l'état permanent

Le passage de votre système de la phase pilote à la mise en œuvre générale est un moment crucial pour faire le point. Travaillez avec une équipe multidisciplinaire pour réfléchir par exemple à ce qui a bien fonctionné, à ce qui n'a pas fonctionné et à ce que vous changeriez si vous aviez su alors ce que vous savez maintenant.

- Avez-vous déterminé que le passage à l'état permanent nécessite des mesures de protection, une mobilisation ou une divulgation supplémentaires?
- Votre projet nécessitera-t-il des logiciels ou du matériel supplémentaires?
- Avez-vous mis le projet à l'essai pour en tester l'extensibilité?
- Avez-vous tiré des leçons au cours de la phase pilote qui pourraient orienter l'utilisation future par IRCC du soutien automatisé à la prise de décision?

L'établissement d'un rapport officiel à la fin du projet ne constitue pas seulement une bonne pratique qui permet d'assurer une gestion efficace du projet. Un tel rapport renforce également la responsabilité et la transparence, et permet davantage aux autres de tirer des leçons de votre expérience.

Veuillez inscrire vos observations, vos analyses et des liens vers des documents clés ici.

PARTIE III – APERÇU DES CONSIDÉRATIONS JURIDIQUES ET CONSEILS PRATIQUES

La notion de règle de droit aide à garantir que la loi est appliquée de manière transparente et prévisible, et impose une forme de responsabilité aux personnes touchées par les décisions administratives. Les mêmes principes juridiques et obligations de droit public s'appliquent lorsque l'IA est utilisée pour faciliter la prise de décisions administratives.

Vous trouverez ci-dessous quelques considérations juridiques clés – concernant notamment le droit administratif, les droits de la personne et la protection de la vie privée – qui se posent lors de l'utilisation de l'IA et de l'automatisation pour soutenir la prise de décisions administratives. Ces considérations sont destinées à guider le personnel d'IRCC tout au long du cycle de vie de l'IA et sont accompagnées d'un ensemble de conseils pratiques pour aider à atténuer les risques.

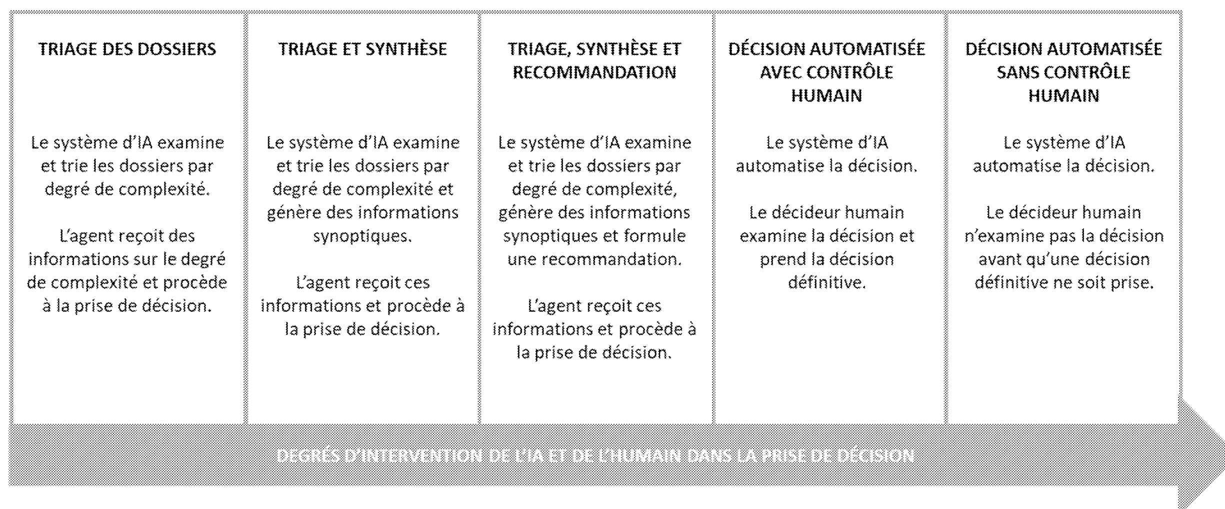
Pour vous aider à relever les risques juridiques et à adopter des stratégies d'atténuation pertinentes tout au long des différentes étapes de votre projet, vous pourriez également travailler avec divers experts multidisciplinaires – y compris ceux des directions générales chargées des politiques et des programmes, des opérations, des données, de la protection de la vie privée, de la gestion des risques, de l'expérience client, des communications et de la transformation d'IRCC. Étant donné que les considérations juridiques varient selon le contexte, **vous devriez également consulter régulièrement l'Unité des services juridiques d'IRCC** pour obtenir des conseils concernant les cas d'utilisation en particulier (c'est également obligatoire selon la *Directive sur la prise de décision automatisée* du SCT).

Droit administratif

Les responsabilités en matière de droit administratif se concentrent sur trois concepts essentiels : la légalité, l'équité et le caractère raisonnable de la prise de décision. Les questions qui nécessitent une attention particulière dans l'environnement d'IRCC sont les suivantes :

- Garantir l'équité procédurale – une participation significative est un élément essentiel de l'équité procédurale. Cela veut dire que le demandeur connaît à l'avance les critères législatifs et les informations factuelles utilisés pour la prise de décision et a la possibilité d'agir en conséquence;
- Garantir un processus décisionnel défendable – les décisions sont raisonnables et leur justification suffisamment transparente et compréhensible;
- Éviter de se fier indûment ou de manière inappropriée à l'IA – les agents comprennent la portée prévue d'un système d'IA et utilisent les informations fournies par le système en conséquence;
- Prendre en compte les critères et les informations appropriés, et atténuer les biais – les informations utilisées par un système d'IA sont conformes aux critères législatifs.

Un certain nombre de facteurs, notamment le degré d'intervention humaine dans la conception et le développement d'un système d'IA et les motifs des changements apportés au système au fil du temps, peuvent avoir un impact sur les considérations de droit administratif. Cependant, en général, plus un outil d'IA guide un processus décisionnel, plus il faut veiller à ce que les principes du droit administratif continuent d'être respectés. Le gouvernement module notamment l'ampleur des mesures d'atténuation des risques posés par les systèmes automatisés en fonction de l'importance des risques grâce à son outil d'évaluation de l'incidence algorithmique et à sa *Directive sur la prise de décision automatisée*.



Droits de la personne

Le droit des droits de la personne fournit une autre lentille importante pour réfléchir aux considérations juridiques qui se posent lorsqu'on utilise l'IA pour soutenir la prise de décision. Il est pertinent en raison de la nature de certaines décisions administratives et du fait que les outils d'IA pourraient conduire de façon imprévue à des décisions soulevant des considérations relatives aux droits de la personne. Par exemple, un système d'IA pourrait utiliser par inadvertance l'adresse ou les informations de localisation géographique d'un client comme un indicateur de l'origine nationale ou ethnique, ou de l'appartenance à un groupe religieux particulier. Dans le contexte d'IRCC, il est particulièrement important de protéger les droits de la personne par la promotion de l'égalité et la lutte contre la discrimination.

Lors de l'élaboration et du déploiement d'un modèle d'IA, il faut se poser les deux questions clés suivantes :

1) **Le modèle crée-t-il des distinctions fondées sur des motifs interdits ou analogues?** Les motifs interdits sont explicitement énumérés à l'article 15 de la *Charte canadienne des droits et libertés*, c'est-à-dire la race, l'origine nationale ou ethnique, la couleur, la religion, le sexe, l'âge ou les déficiences intellectuelles ou physiques. Les motifs analogues sont des motifs de discrimination interdits identifiés par les tribunaux, par exemple la non-citoyenneté, l'état civil, l'orientation sexuelle, le statut de personne adoptée et le lieu de naissance.

2) **Ces distinctions sont-elles discriminatoires?** Veuillez noter que les différences de traitement ne sont pas toutes discriminatoires. Dans certains cas, il peut être pertinent de prendre en compte des caractéristiques personnelles à la lumière de critères législatifs.

Droit relatif au respect de la vie privée

En vertu de la *Loi sur la protection des renseignements personnels*, la responsabilité des institutions gouvernementales est axée sur le respect des principes et des obligations liés à la collecte, à la notification, à l'utilisation, à l'exactitude, à la divulgation, à la conservation et à l'élimination des renseignements personnels. Bien que la Loi soit conçue pour être neutre sur le plan technologique, le

risque d'intrusion dans la vie privée pourrait être plus important dans le contexte de l'IA et de l'automatisation en raison du volume important d'informations personnelles traitées.

Le personnel d'IRCC qui participe aux projets d'IA doit pratiquer une gestion rigoureuse des données et des dossiers, et prendre des mesures pour s'assurer que les outils d'IA n'utilisent les renseignements personnels recueillis qu'aux fins initialement prévues. Familiarisez-vous avec les règles et exigences applicables en matière de protection de la vie privée et consultez des experts en la matière dès le début des projets d'IA pour vous assurer que les considérations relatives à la protection de la vie privée restent au premier plan pendant tout le cycle de vie de l'IA.

Conseils pratiques pour se conformer au droit administratif, aux droits de la personne et au droit relatif au respect de la vie privée

Vous trouverez ci-après des conseils pratiques d'étapes à suivre pour « mettre la loi en pratique » lors de l'élaboration de projets d'IA. Ces conseils sont destinés à aider les praticiens à passer de l'interprétation et de l'application des cadres juridiques à l'identification des principaux risques dans les domaines du droit administratif, des droits de la personne et du droit relatif au respect de la vie privée, et à atténuer ces risques.

- **Déterminer quels renseignements personnels peuvent être utilisés pour atteindre un but quelconque et n'utiliser que ceux qui portent le moins atteinte à la protection de la vie privée.** Vous devriez conserver les renseignements sur les essais de modélisation et sur tout compromis retenu pour justifier les décisions touchant l'utilisation des renseignements personnels.
- **S'assurer du respect de l'obligation de notification de collecte de nouveaux renseignements personnels ou d'utilisations nouvelles de renseignements personnels recueillis précédemment.** IRCC peut réutiliser les renseignements personnels à des fins similaires à celles pour lesquels ils ont été recueillis initialement ou avec le consentement des personnes concernées. Toutefois, il faut effectuer un examen minutieux pour établir un lien suffisamment direct entre le but de l'utilisation initiale des renseignements personnels et le but proposé de l'utilisation de ces renseignements par un système d'IA.
- **S'assurer que les données utilisées soient exactes, à jour, fiables et aussi complètes que possible.** Vous devriez notamment évaluer l'adéquation des données d'apprentissage d'un système d'IA avec le but prévu du système, la fiabilité des données en tant qu'échantillon représentatif et l'exactitude des données personnelles dans le jeu de données pendant le cycle de vie de leur utilisation prévue. En travaillant de concert avec la Direction générale du dirigeant principal des données d'IRCC et les responsables des programmes et politiques concernés, vous pourrez vous assurer que les données historiques sont adaptées à des fins telles que l'apprentissage d'un modèle d'IA.
- **Évaluer dans quelle mesure on peut anonymiser[2] les données.** Bien qu'efficace pour réduire au minimum l'intrusion dans la vie privée, l'anonymisation des données peut être difficile à

[2] L'anonymisation des données consiste à modifier les données de telle sorte qu'elles ne puissent plus être attribuées à une personne identifiable et qu'elles ne puissent pas être réorganisées pour identifier à nouveau une personne.

mettre en œuvre, car il pourrait devenir possible de relier des informations à une personne identifiable à mesure que de nouvelles informations sont recueillies et que les moyens techniques pour empêcher la dépersonnalisation deviennent plus sophistiqués.

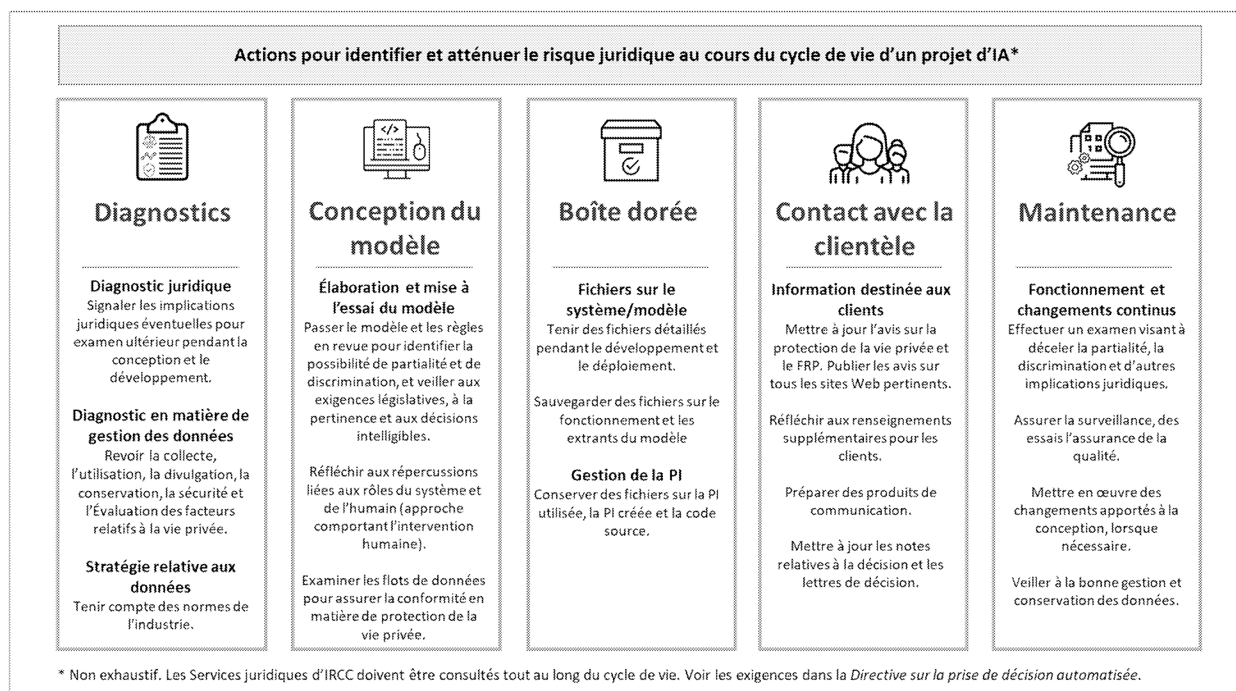
- **Évaluer les flux de données pour s'assurer que la collecte, l'utilisation et la divulgation des renseignements personnels répondent aux exigences juridiques et politiques.** Les concepteurs de modèles et les experts en données s'assureront de bien comprendre le fonctionnement du modèle et les données qu'il utilise, tandis que les experts en protection de la vie privée pourront vous conseiller sur la nécessité ou non de réaliser une évaluation des incidences sur la vie privée.
- **Évaluer les données d'apprentissage pour repérer les biais et évaluer si le modèle ou les règles peuvent les détecter et les corriger, ou potentiellement les intégrer ou les exacerber.** Cette évaluation doit être effectuée dès les premières étapes de la conception et de façon continue, car les biais peuvent s'infiltrer involontairement dans un modèle au fil du temps. On peut tirer parti de l'IA pour détecter et atténuer les préjugés et les discriminations cognitives individuelles, afin d'améliorer la qualité et la cohérence de la prise de décision.
- **Tenir compte des différences de fonctionnalité avec les modèles d'IA et déterminer si elles ont un impact sur les obligations de confidentialité.** On pourrait notamment évaluer à quel point différents modèles ou différentes solutions à l'étude présentent un risque global de porter atteinte à la vie privée.
- **Évaluer l'*algorithme* et son adéquation à l'utilisation prévue, notamment en se fondant sur les meilleures pratiques, les normes du secteur (le cas échéant) et toute autre règle ou réglementation applicable.** En plus de consulter des partenaires internes, vous devrez peut-être mener un processus d'évaluation par les pairs formel ou informel auprès d'intervenants externes.
- **Évaluer le modèle d'IA, les règles d'inclusion dans l'IA, les critères législatifs et d'autres informations pertinentes.** Cette évaluation devrait être réalisée avec la participation d'un agent expérimenté qui pourrait s'assurer que le modèle et les règles sont conformes aux processus décisionnels habituels.
- **Déterminer si les variables d'entrée ou les règles du modèle d'IA tiennent compte des motifs interdits ou analogues de discrimination, directement, indirectement ou par l'entremise d'un tiers.** Pour le déterminer, il faut bien comprendre les données et le contexte, et collaborer avec les concepteurs du modèle, les experts en données et les conseillers techniques.
- **Déterminer s'il est pertinent et approprié de prendre en compte les motifs interdits ou analogues de discrimination, par exemple à la lumière des critères législatifs ou de leur pertinence pour la prise de décision.** Les conseillers techniques, les responsables des politiques/programmes et les agents expérimentés peuvent apporter leur expertise pour faire des choix éclairés à cet égard. Par exemple, les critères législatifs pour devenir un immigrant économique tiennent compte de l'âge du demandeur, puisque lorsque l'âge est pris en considération avec d'autres facteurs, il est un indicateur pertinent de choses telles que la durée pendant laquelle le demandeur est susceptible de rester sur le marché du travail et sa capacité à s'adapter à la vie au Canada.
- **Rechercher les tendances qui peuvent suggérer un traitement différent basé sur des motifs interdits ou analogues de discrimination, et évaluer si cette différence de traitement**

pourrait entraîner des conséquences discriminatoires. Cette évaluation doit être faite en consultation avec l'équipe du droit des technologies de l'information d'IRCC, des agents expérimentés et d'autres experts multidisciplinaires. Vous devrez peut-être procéder à une analyse plus approfondie pour déterminer la cause de traitements discriminatoires (p. ex. données de formation, algorithme, influence humaine, conception de règles) et prendre des mesures pour résoudre le problème.

- **Évaluer la technologie et son utilisation par les décideurs avec l'aide d'experts multidisciplinaires afin de répondre aux exigences de transparence et d'explicabilité.** On peut consigner les motifs d'une décision de deux manières : créer et conserver une documentation de la décision concernant chaque demande, ou créer la possibilité de reconstruire les étapes de traitement et les motifs de la décision après coup.
- **Évaluer l'explicabilité et la pertinence des décisions et des motifs disponibles.** En fonction du contexte, les décisions qui s'inscrivent dans la fourchette des résultats acceptables sont défendables – par exemple, éviter de se fier indûment à l'IA, restreindre le pouvoir discrétionnaire, éviter les préjugés, appliquer les critères législatifs, prendre en compte de manière appropriée les informations pertinentes et les raisons convaincantes. Les décideurs peuvent également utiliser l'IA pour les aider à motiver leurs décisions de manière à ce qu'elles soient plus facilement comprises par les personnes concernées, ce qui renforce la responsabilisation et facilite éventuellement l'accès à un contrôle judiciaire.
- **Évaluer les exigences en matière de tenue de registres.** Les renseignements utilisés à des fins de décision administrative doivent être conservés pendant au moins deux ans. On devrait également envisager de conserver des renseignements sur les mesures qui ont été prises pour assurer la protection de la vie privée dès la conception et le déploiement des outils d'IA.
- **Fournir des informations et une formation adéquates aux décideurs** pour s'assurer qu'ils comprennent la nature de l'outil d'IA et la répartition des rôles entre le décideur humain et l'outil d'IA.
- **Surveiller le fonctionnement du système d'IA en permanence** pour détecter et éviter les risques liés au droit administratif, aux droits de la personne et à la protection de la vie privée. Cette surveillance peut inclure une analyse globale pour détecter tout impact discriminatoire restant, ou le suivi des risques de non-respect du droit administratif ou du droit relatif à la vie privée qui peuvent apparaître au fil du temps.

S'engager dans le cycle de vie du projet d'IA

Le tableau ci-dessous donne un aperçu général des étapes à suivre tout au long du cycle de vie d'un projet d'IA – depuis l'élaboration de diagnostics et la conception d'un modèle approprié dès le début du projet jusqu'à la notification adéquate des clients et la garantie d'un suivi continu. La gestion et la conservation des dossiers – c'est-à-dire la « boîte dorée » – sont également essentielles à une utilisation responsable de l'IA à tous les stades du cycle de vie. Il peut s'agir de dossiers de consultation, de règles types et de modifications de modèles, d'analyses statistiques, d'assurance qualité, de décisions clés concernant la conception et le déploiement, d'instructions fournies aux agents et de mesures prises pour atténuer le risque juridique en permanence.


Actions pour identifier et atténuer le risque juridique au cours du cycle de vie d'un projet d'IA*
Diagnostics
Diagnostic juridique
Signaler les implications juridiques éventuelles pour examen ultérieur pendant la conception et le développement.
Diagnostic en matière de gestion des données
Revoir la collecte, l'utilisation, la divulgation, la conservation, la sécurité et l'Évaluation des facteurs relatifs à la vie privée.
Stratégie relative aux données
Tenir compte des normes de l'industrie.
Conception du modèle
Élaboration et mise à l'essai du modèle
Passer le modèle et les règles en revue pour identifier la possibilité de partialité et de discrimination, et veiller aux exigences législatives, à la pertinence et aux décisions intelligibles.
Réfléchir aux répercussions liées aux rôles du système et de l'humain (approche comportant l'intervention humaine).
Examiner les flots de données pour assurer la conformité en matière de protection de la vie privée.
Boîte dorée
Fichiers sur le système/modèle
Tenir des fichiers détaillés pendant le développement et le déploiement.
Sauvegarder des fichiers sur le fonctionnement et les extrants du modèle
Gestion de la PI
Conserver des fichiers sur la PI utilisée, la PI créée et la code source.
Contact avec la clientèle
Information destinée aux clients
Mettre à jour l'avis sur la protection de la vie privée et le FRP. Publier les avis sur tous les sites Web pertinents.
Réfléchir aux renseignements supplémentaires pour les clients.
Préparer des produits de communication.
Mettre à jour les notes relatives à la décision et les lettres de décision.
Maintenance
Fonctionnement et changements continus
Effectuer un examen visant à déceler la partialité, la discrimination et d'autres implications juridiques.
Assurer la surveillance, des essais l'assurance de la qualité.
Mettre en œuvre des changements apportés à la conception, lorsque nécessaire.
Veiller à la bonne gestion et conservation des données.
* Non exhaustif. Les Services juridiques d'IRCC doivent être consultés tout au long du cycle de vie. Voir les exigences dans la Directive sur la prise de décision automatisée.

PARTIE IV – EXIGENCES DE BASE EN MATIÈRE DE PROTECTION DE LA VIE PRIVÉE

Au Canada, on estime que la protection de la vie privée est un droit de la personne. Compte tenu du fait que la majorité des données que nous traitons à Immigration, Réfugiés et Citoyenneté Canada (IRCC) sont des renseignements personnels, il importe de se rappeler les exigences en matière de protection de la vie privée au moment de planifier, d'élaborer et de surveiller toute initiative exploitant une technologie axée sur les données. Les incidences et les nuances liées à la protection de la vie privée varieront selon les détails de l'initiative, mais certaines exigences de base doivent être respectées pour toutes les initiatives.

Le présent document vise à aider les secteurs de programme et les responsables de l'élaboration des politiques à tenir compte de la protection de la vie privée dès le début d'une initiative de technologie perturbatrice et à aider les responsables de programme à mettre en œuvre ou à modifier les procédures, les processus et la gouvernance, s'il y a lieu.

Tout au long du document, on utilise l'expression « technologie perturbatrice » pour englober les techniques de science des données et les technologies axées sur les données telles que l'analyse avancée, l'apprentissage machine, l'intelligence artificielle et l'automatisation. Une technologie perturbatrice est une nouvelle technologie qui a une incidence rapide et importante sur les technologies qui existaient auparavant[3]. En outre, le document doit être lu en parallèle avec le Guide de politique sur le soutien automatisé sur la prise de décision.

Tout au long du document, vous trouverez des exigences relatives à la protection de la vie privée qui s'appliquent quel que soit le type de technologie perturbatrice ou d'automatisation envisagé. Il s'agit d'**exigences minimales**, et il est possible de prévoir d'autres mesures de protection de la vie privée. Les exigences en matière de protection de la vie privée énoncées dans le présent document sont fondées sur la *Loi sur la protection des renseignements personnels*, les politiques, directives et lignes directrices du Secrétariat du Conseil du Trésor, et les orientations internes d'IRCC.

Le respect des exigences énoncées ci-après ne remplace **pas** une évaluation officielle des facteurs relatifs à la vie privée (EFVP). Ces exigences aideront plutôt les secteurs de programme à déterminer ce qu'ils doivent faire pour se conformer aux exigences en matière de protection de la vie privée, et peuvent contribuer à la réalisation des évaluations de la protection de la vie privée, le cas échéant. En ce qui a trait aux programmes dans lesquels on a l'intention d'intégrer une technologie perturbatrice, il faudra remplir une Évaluation des besoins en matière de protection de la vie privée (lien interne) et l'envoyer à la Division de l'AIPRP : ATIPinternal-AIPRPinterne@cic.gc.ca.

[3] Définition adaptée du Grand dictionnaire terminologique http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8355620. Consulté le 15 février 2021.

Fondement législatif

Un programme doit déterminer les pouvoirs conférés par le Parlement en ce qui a trait à la collecte et à l'utilisation des renseignements personnels aux fins précises du programme.

À IRCC, la Partie 4.1 de la *Loi sur l'immigration et la protection des réfugiés* permet aux programmes qui relèvent de la LIPR d'utiliser des moyens électroniques pour administrer la *Loi*, et le paragraphe 2.2 du *Décret sur les passeports canadiens* confère ce pouvoir au Programme des passeports. Si vous avez des doutes concernant votre pouvoir légal de recueillir et d'utiliser des renseignements personnels dans le cadre d'une initiative axée sur une technologie perturbatrice, veuillez communiquer avec les Services juridiques.

Collecte de renseignements

Seuls les renseignements personnels qui ont un lien direct avec un programme ou une activité d'IRCC peuvent être recueillis, et chaque élément de données des renseignements personnels doit être nécessaire à l'administration du programme[4]. Chaque fois que c'est possible, les renseignements personnels seront recueillis auprès de l'individu lui-même[5].

Aux fins des initiatives exploitant une technologie perturbatrice, **seuls les renseignements trouvés dans les systèmes d'enregistrement ministériels (p. ex. le Système mondial de gestion des cas ou SMGC) devraient être utilisés**. Par ailleurs, les activités des modèles et algorithmes d'entraînement devraient être effectués en dehors de ces systèmes d'enregistrement. On ne doit pas utiliser les données recueillies auprès de sources extérieures, à moins que leur nécessité ne soit démontrée, et des accords d'échange d'information, des protocoles d'entente ou des accords sur les niveaux de service, notamment, doivent être mis en place en relation avec cette utilisation et respectés

Lorsque vous songez à traiter des renseignements personnels à l'aide d'un modèle ou d'un algorithme, posez-vous les questions suivantes[6] :

- nécessité démontrée : y a-t-il une nécessité clairement définie pour la collecte et l'utilisation des renseignements personnels? Quel problème devrait-on résoudre grâce à ces renseignements?
- proportionnalité : la collecte et l'utilisation des renseignements personnels sont-elles suffisamment ciblées et personnalisées afin d'être raisonnablement proportionnelles à la réduction du droit à la vie privée de l'individu que la collection peut représenter?
- efficacité : les renseignements personnels sont-ils efficaces en vue de régler le problème et sont-ils clairement associés à la résolution du problème?

[4] Article 4 de la *Loi sur la protection des renseignements personnels.*
[5] Paragraphe 5(1) de la *Loi sur la protection des renseignements personnels.*
[6] Adaptation de la publication du Commissariat à la protection de la vie privée intitulée *Une question de confiance : Intégrer le droit à la vie privée aux mesures de sécurité publique au 21e siècle* : https://www.priv.gc.ca/fr/sujets-lies-a-la-protection-de-la-vie-privee/surveillance/police-et-securite-publique/gd_sec_201011/#toc5a. Consulté le 29 janvier 2021.

- atteinte à la vie privée minimale : les éléments de données spécifiques choisis sont-ils les renseignements disponibles les moins envahissants? Assurez-vous que toutes les autres options d'enquête moins envahissantes ont été épuisées.
- exactitude et validité : la source de données est-elle une source d'information fiable et valable? Comment allez-vous vous assurer que les données sont exactes et à jour?

Avis et information quant aux fins de la collecte

IRCC doit informer les personnes des fins pour lesquelles leurs renseignements sont recueillis, ce qu'on appelle communément un « avis de confidentialité »[7]. Cet avis doit être donné au moment de la collecte ou avant.

Selon la technologie utilisée**, il pourrait être nécessaire de mettre à jour les avis de confidentialité** et les fichiers de renseignements personnels (FRP) actuels pour tenir compte de l'utilisation de la technologie. Si les renseignements personnels doivent être utilisés ou communiqués dans un but autre que la fin originale ou l'usage compatible, un avis officiel de confidentialité sera probablement exigé.

Transparence

IRCC doit informer les anciens demandeurs que leurs renseignements ont été utilisés pour entraîner ou construire des modèles.

La mise à jour du site Web public d'IRCC sur l'analyse et l'automatisation est un moyen suggéré pour informer le public que les demandes antérieures ont influencé les modèles actuels. Un avis en ligne pouvant être modifié au fur et à mesure qu'IRCC modifie son utilisation de la technologie peut également aider à satisfaire les exigences de transparence.

Explicabilité

Les personnes ont le droit de savoir exactement comment leurs renseignements personnels ont été traités par un système de technologie perturbatrice. En veillant à offrir sur demande des explications en langage simple, les personnes pourraient voir comment la technologie a été utilisée pour appuyer la prise de décision.

[7] Paragraphe 5(2) de la *Loi sur la protection des renseignements personnels.*

Exactitude

IRCC doit veiller, dans la mesure du possible, à ce que les renseignements personnels qu'il utilise à des fins administratives soient à jour, exacts et complets[8]. Il doit également veiller à ce que des mécanismes soient mis en place pour corriger les renseignements inexacts.

Pour ce qui est des initiatives associées à une technologie perturbatrice, il faut notamment s'assurer de recueillir les données à partir d'une source fiable, de veiller à la qualité des données, d'élaborer des mécanismes technologiques garantissant le fonctionnement de la technologie (comme les boucles de rétroaction et les tests en aveugle) et d'assurer la qualité des résultats. En outre, pour garantir l'exactitude des données, les secteurs de programme doivent prendre les mesures nécessaires pour minimiser les biais involontaires dans les données. Enfin, l'exactitude suppose la tenue à jour du modèle et la garantie que le modèle est entraîné et recyclé selon les données les plus récentes, les plus précises et les plus fiables.

Utilisation

Les renseignements personnels ne doivent être utilisés qu'aux fins pour lesquelles ils ont été initialement recueillis, pour les usages qui sont compatibles avec ces fins ou aux fins pour lesquelles ils peuvent être communiqués en vertu du paragraphe 8(2) (voir la section « Communication » ci-dessous)[9].

L'application d'une technologie perturbatrice à un ensemble de données comportant des renseignements personnels représente un usage, peu importe qu'il soit de nature administrative (dans le cadre d'un processus décisionnel qui touche la personne) ou non. Les renseignements personnels doivent être traités de manière appropriée, quel que soit le niveau d'automatisation ou de soutien que la technologie fournit. Vous devez également vous assurer que l'utilisation d'une technologie perturbatrice présente un usage compatible des renseignements personnels. Pour déterminer ce qui constitue un usage compatible des renseignements personnels, les fins originales et les fins proposées doivent être si étroitement liées que la personne s'attendrait à ce que les renseignements soient utilisés pour des fins compatibles, même si l'usage n'est pas précisé. Dans le cas des programmes dont les responsables ne savent pas exactement s'il s'agit d'un usage compatible des renseignements personnels, ces derniers doivent consulter les Services juridiques.

Les éléments suivants peuvent contribuer à réduire la probabilité et l'incidence d'un usage inapproprié.

- **Collecte des renseignements nécessaires uniquement** : n'utilisez que les éléments de données personnelles absolument nécessaires. Par exemple, plutôt que le nom complet et l'identificateur-client unique (ICU), seul l'ICU est exigé, ou plutôt que l'un ou l'autre, un numéro d'identification unique peut être généré pour les besoins du modèle uniquement.
- **Réduction de la granularité des données** : il s'agit de rendre moins précis certains éléments de données lorsqu'il n'est pas essentiel d'avoir des renseignements précis. Par exemple, la date de

[8] Paragraphe 6(2) de la *Loi sur la protection des renseignements personnels*.
[9] Article 7 de la *Loi sur la protection des renseignements personnels*.

naissance au complet peut ne pas être requise, seule l'année de naissance suffit et produit des résultats efficaces, ou le numéro de téléphone complet peut ne pas être nécessaire et seul suffit l'indicatif régional.

- **Dépersonnalisation** : il peut être possible de dépersonnaliser la totalité ou une partie des ensembles de données afin de réduire la probabilité que des personnes soient identifiées et de réduire les risques d'atteinte à la vie privée. Les ensembles de données peuvent même être dépersonnalisés dans la mesure où il ne s'agit plus de renseignements personnels, ce qui permettrait de diminuer les restrictions liées aux données, tout en produisant des résultats utiles.
- **Besoin de savoir** : seules les personnes qui ont besoin de connaître les renseignements devraient y avoir accès, quel que soit le type de technologie perturbatrice appliquée. Par exemple, on peut créer différentes versions des résultats : une version détaillée pour les agents qui présente un degré élevé de renseignements personnels, et une version résumée pour la direction qui contient moins de renseignements personnels. Il peut aussi s'agir de limiter le nombre de personnes qui ont accès aux données d'entraînement et à l'endroit où les résultats sont enregistrés. Enregistrer les données d'une manière permettant de modifier et d'adapter les droits d'accès peut aider en ce sens.

Communication

Les renseignements personnels que détient IRCC ne doivent être communiqués à personne ni à aucune organisation pour une raison ou une autre, sauf pour les raisons énumérées au paragraphe 8(2) de la *Loi sur la protection des renseignements personnels*[10].

Dans le cas des initiatives associées à une technologie perturbatrice, cela comprend les renseignements trouvés dans les systèmes d'enregistrement ministériels (p. ex. SMGC) que d'autres organismes tels que l'Agence des services frontaliers du Canada (ASFC) ou le Service canadien du renseignement de sécurité (SCRS) peuvent consulter. L'échange régulier de renseignements peut se poursuivre entre IRCC et les organismes partenaires, mais les protocoles d'entente (PE), les ententes sur l'échange de renseignements (EER) et d'autres accords officiels doivent être mis à jour et modifiés par les voies appropriées si l'on souhaite communiquer régulièrement des résultats. Les groupes de l'Admissibilité et des Relations internationales et intergouvernementales (lien interne) d'IRCC peuvent fournir des orientations.

Mesures de protection

Les renseignements personnels doivent être protégés de manière appropriée, quel que soit le type de technologie qui leur est appliqué.

Des mesures de protection administratives, techniques et physiques appropriées devraient être appliquées aux renseignements personnels pour chaque programme utilisant des s'étendent donc à toutes les étapes d'une initiative liée à une technologie perturbatrice. Comme les autres efforts et opérations du programme, et il faudrait envisager de réduire la probabilité d'atteinte à la vie privée et à la sécurité tout au long du développement de nouvelles solutions technologiques. Les programmes ont la

[10] Article 8 de la *Loi sur la protection des renseignements personnels.*

responsabilité de mener à bien le processus d'évaluation et d'autorisation de Sécurité des TI (lien interne). En outre, les programmes ont la responsabilité d'élaborer des mesures visant à réduire les répercussions en cas d'atteinte à la vie privée ou à la sécurité. Les spécialistes de la Sécurité des TI (lien interne) peuvent fournir des orientations plus détaillées.

Conservation et retrait

Les renseignements personnels utilisés à des fins administratives (dans le cadre d'un processus décisionnel qui touche la personne) doivent être conservés pendant au moins deux ans, et conformément aux calendriers de conservation et de disposition[11] (lien interne). Toutes les données (sauf les données d'entraînement dans la zone d'exploration) doivent être conservées de telle manière qu'en cas de plainte ou de poursuite la décision puisse être reproduite.

En ce qui a trait aux initiatives reposant sur une technologie perturbatrice, les décisions concernant la conservation et le retrait seront fondées sur les aspects suivants :

- données d'entraînement : devraient être détruites ou retirées de la zone d'exploration lorsqu'elles ne sont plus requises pour les besoins d'entraînement;
- modèle ou algorithme : devrait faire l'objet d'un contrôle des versions de la même manière que les pratiques de conservation des logiciels;
- résultats : selon le type de modèle, les résultats devraient être ajoutés à un calendrier de conservation et de retrait existant, un nouveau calendrier devrait être créé, ou les résultats devraient être considérés comme transitoires, et détruits lorsqu'ils ne sont plus nécessaires.

La Direction de la gestion de l'information (lien interne) d'IRCC peut fournir des conseils sur la conservation et le retrait des documents créés dans le cadre de l'élaboration et du déploiement d'une technologie perturbatrice.

Accès aux renseignements personnels

Les personnes ont le droit d'accéder aux renseignements personnels qu'IRCC détient à leur sujet[12]. IRCC devrait toujours disposer de moyens pour communiquer des renseignements personnels sur demande.

Les documents relatifs à l'élaboration, au déploiement et à la surveillance des initiatives axées sur une technologie perturbatrice devraient aussi être accessibles en cas d'une demande au titre de la *Loi sur l'accès à l'information* ou de la *Loi sur la protection des renseignements personnels*.

[11] Paragraphe 6(1) de la *Loi sur la protection des renseignements personnels.*
[12] Paragraphe 12(1) de la *Loi sur la protection des renseignements personnels.*

Plaintes et examens

> Conservez des documents clairs et accessibles qui décrivent toutes les mesures et décisions prises pour assurer le respect de la vie privée. Ils devront être produits en cas de plainte ou d'examen.

Les personnes ont le droit de déposer une plainte officielle auprès du Commissariat à la protection de la vie privée du Canada (CPVP) sur la façon dont IRCC traite et gère leurs renseignements personnels, y compris la façon dont ces renseignements sont traités par la technologie[13]. Le CPVP peut également effectuer des examens de programmes lorsqu'il le juge nécessaire, et il a le pouvoir d'obtenir toute information relative au traitement des renseignements personnels dans le cadre du programme, y compris des détails sur les modèles, les algorithmes et l'automatisation.

Surveillance

> La surveillance du respect de la protection de la vie privée par rapport aux exigences susmentionnées doit être intégrée dès l'étape de l'élaboration du modèle, et un calendrier de surveillance devrait être établi après le déploiement qui implique les propriétaires du programme ainsi que les architectes du modèle.

En plus de veiller à ce que la technologie perturbatrice fonctionne correctement, voici une liste non exhaustive des activités de surveillance à prévoir.

- **Collecte** : s'assurer qu'aucune donnée provenant de sources autres que les systèmes d'enregistrement ministériels (p. ex. SMGC) et autres dépôts de données d'IRCC n'est incluse dans la technologie perturbatrice, et s'il y a des données de l'extérieur, veiller à ce que l'EER ou le protocole d'entente approprié soit en place et à jour.
- **Avis** : examiner les avis de confidentialité et les pratiques de transparence et d'explicabilité pour garantir l'exactitude et la mise à jour des informations. Les mettre à jour chaque fois que cela est nécessaire.
- **Conservation et élimination** : revoir les pratiques de conservation et d'élimination et veiller à ce qu'aucune information ne soit conservée plus longtemps qu'elle ne devrait l'être.
- **Exactitude** : intégrer des pratiques courantes en matière de qualité des données afin de garantir que les données sont exactes, à jour et aussi complètes que possible, et modifier les informations au besoin.
- **Utilisation** : revoir les pratiques de traitement des données, s'assurer que les mesures d'atténuation contre les usages inappropriés fonctionnent correctement, employer de nouvelles mesures au besoin, et mettre à jour les pratiques périodiquement.
- **Communication** : revoir les pratiques de communication pour s'assurer que la communication de renseignements se fait dans le cadre de protocoles d'entente ou d'EER à jour, et veiller à ce que les autres organismes gouvernementaux ne puissent voir que les renseignements qu'ils ont l'autorisation de voir dans le SMGC. Apporter les modifications nécessaires.

[13] Article 29 de la *Loi sur la protection des renseignements personnels.*

- **Mesures de protection** : effectuer des contrôles de sécurité réguliers pour s'assurer que les données sur la formation, la technologie elle-même et les résultats sont sûrs.

Ce document donne un aperçu des exigences de base pour la protection de la vie privée que toutes les initiatives technologiques perturbatrices à IRCC doivent prendre en compte. Des exigences supplémentaires peuvent être requises en fonction d'initiatives spécifiques, et même si vous suivez ces exigences, une évaluation de la protection de la vie privée peut être nécessaire. Une approche multipartite est le meilleur moyen de garantir que toutes les exigences et tous les risques sont traités et atténués de manière appropriée. L'équipe de l'AIPRP et d'autres groupes mentionnés dans ce document sont là pour vous aider à naviguer dans les considérations relatives à la protection de la vie privée afin de vous assurer que les programmes sont élaborés en tenant compte de la protection de la vie privée. Veuillez remplir une évaluation des besoins en matière de la vie privée (lien interne) pour commencer votre consultation avec la Division de l'AIPRP.

PARTIE V – LISTE DE VÉRIFICATION POUR LA DIRECTIVE SUR LA PRISE DE DÉCISION AUTOMATISÉE

La *Directive sur la prise de décision automatisée* est la politique du gouvernement du Canada qui régit l'utilisation de l'intelligence artificielle à l'appui des décisions administratives. Cette Directive est entrée en vigueur en avril 2020 et s'applique aux systèmes, aux outils et aux modèles automatisés utilisés par les institutions fédérales pour prendre ou appuyer les décisions qui concernent des clients externes.

La liste de contrôle suivante donne un aperçu des exigences auxquelles vous devez satisfaire à différentes étapes pour vous conformer à la Directive. Vous devriez les considérer comme des exigences minimales; d'autres exigences propres au cadre juridique, aux politiques et aux programmes d'IRCC pourraient également s'appliquer.

Certaines exigences s'appliquent à tous les systèmes visés par la portée de la Directive, tandis que d'autres varient en fonction du niveau d'impact associé à l'algorithme utilisé. De plus, bien que la Directive ne s'applique qu'aux systèmes automatisés en production (excluant donc ceux qui sont au stade de mise à l'essai), l'examen des exigences et de l'outil obligatoire d'évaluation de l'incidence algorithmique au début du processus vous permettra d'anticiper et d'élaborer des stratégies d'atténuation des risques critiques lors de la phase de conception et de mise à l'essai de votre projet.

A. Explorer le soutien automatisé à la prise de décision à titre de solution

- ☐ **Services juridiques** – Avez-vous consulté les Services juridiques d'IRCC? Cette consultation doit se faire dès la conceptualisation et se poursuivre de façon continue.

B. Concevoir votre système

- ☐ **Évaluation de l'incidence algorithmique (EIA)** – Avez-vous effectué une EIA préliminaire et ciblé les exigences associées au niveau d'impact pertinent?
- ☐ **Accès aux composants** – Le cas échéant, avez-vous déterminé la licence appropriée pour les composants logiciels et, si vous utilisez une licence propriétaire, respecté les exigences connexes?
- ☐ **Qualité des données** – Avez-vous confirmé que toutes les données recueillies pour le système décisionnel automatisé et utilisées par celui-ci sont pertinentes, exactes, à jour, et conformes à la Politique sur la gestion de l'information et à la Loi sur la protection des renseignements personnels?
- ☐ **Sécurité** – Avez-vous mené des évaluations des risques liés au système et mis sur pied des mesures de protection appropriées, conformément à la Politique sur la sécurité du gouvernement?
- ☐ **Intervention humaine** – Avez-vous veillé à ce que le système permette une intervention par un humain? (*niveaux III et IV seulement*)

C. Se préparer au lancement

- ☐ **EIA** – Avez-vous répondu de façon intégrale à toutes les questions de l'EIA et appliqué les exigences associées au niveau d'impact pertinent?
- ☐ **EIA** – Avez-vous diffuse les résultats finaux de l'EIA dans un format accessible?

- ☐ **Code source** – Avez-vous rendu public le code source créé sur mesure détenu par le gouvernement du Canada, le cas échéant? Vous devrez aussi déterminer les restrictions d'accès appropriée qui s'applique au code source rendu public.
- ☐ **Évaluer et surveiller les résultats** – Avez-vous élaboré des processus d'évaluation des données et de l'information utilisés dans le système de prise de décisions automatisée visant à s'assurer de l'absence de biais imprévus dans les données et d'autres facteurs qui pourraient influencer injustement les résultats?
- ☐ **Évaluer et surveiller les résultats** – Avez-vous élaboré des processus de surveillance des résultats afin d'éviter les résultats imprévus et de vérifier régulièrement la conformité aux dispositions législatives pertinentes, ainsi qu'à la présente Directive?
- ☐ **Parvenir un avis avant les décisions** – Avez-vous publié, par l'entreprise des sites Web pertinentes, des avis sur le fait que la décision rendue sera formulée totalement ou en partie par un système décisionnel automatisé? Note : ces avis doivent être affichés de façon évidente et fournis en langage simple; les Services juridiques, la division de l'AIPRP et l'équipe des Communications d'IRCC peuvent fournir des conseils à cet effet. (*niveaux II à IV seulement*)
- ☐ **Examen par les pairs** – Avez-vous consulté des experts qualifiés appropriés pour demander une évaluation du système? (*niveaux II à IV seulement*)
- ☐ **Formation des employés** – Avez-vous élaboré la formation appropriée et mis en place des mesures afin que cette formation soit offerte aux employés? (*niveaux II à IV seulement*)
- ☐ **Éventualités** – Avez-vous mis en place des systèmes ou des processus d'urgence? (*niveaux III et IV seulement*)
- ☐ **Approbation** – Avez-vous obtenu le niveau approprié d'approbation pour que votre système soit mis en œuvre? (*Cette exigence s'applique notamment aux systèmes des niveaux III et IV. Cependant, le sous-ministre adjoint chargé du programme qui utilise le système décisionnel automatisé doit veiller à ce que toutes les exigences pertinents de la présente Directive soient respectées.*)

D. Une fois le système fonctionnel

- ☐ **EIA** – Avez-vous mis à jour l'EIA afin qu'elle tienne compte de changements à la fonctionnalité ou à la portée du système? En cas de changement touchant le niveau d'impact, il se peut que vous deviez satisfaire à des exigences supplémentaires.
- ☐ **Explications** – Avez-vous fourni une explication significative aux personnes concernées sur la façon dont la décision a été prise et la raison pour laquelle elle a été prise?
- ☐ **Résultats** – Conservez-vous les résultats des décisions, le cas échéant, et conformez-vous aux exigences en matière d'évaluation et de surveillance?
- ☐ **Évaluer et surveiller les résultats** – Effectuez-vous régulièrement des vérifications des données utilisées par le système pour veiller à ce qu'elles soient toujours pertinentes, exactes et à jour?
- ☐ **Recours** – Avez-vous offert aux clients toute possibilité à leur disposition en matière de recours applicable afin de contester la prise de décision administrative automatisée?
- ☐ **Établissement de rapports** – Avez-vous publié des renseignements sur l'efficacité et l'efficience des systèmes décisionnels automatisés en matière de réalisation des objectifs du programme?

GLOSSAIRE DES TERMES CLÉS

Algorithme : Séquence d'instructions, de règles et de calculs exécutée par un ordinateur selon un ordre particulier pour obtenir un résultat, habituellement une réponse à un problème donné. Il est possible de combiner plusieurs algorithmes pour résoudre des problèmes complexes.

Analytique des données (y compris l'analytique avancée) : Extraction de signification de grands volumes de données (parfois à haute vitesse) au moyen de systèmes informatiques spécialisés. Ces systèmes transforment, organisent et modélisent les données pour tirer des conclusions et cerner des tendances.

Apprentissage machine : Sous-catégorie de l'intelligence artificielle, l'apprentissage machine désigne les algorithmes et les modèles statistiques qui apprennent à partir d'exemples, de données et d'expériences au lieu de suivre des règles préprogrammées. Les systèmes d'apprentissage machine exécutent efficacement une tâche précise sans suivre d'instructions explicites en s'appuyant plutôt sur des modèles et des suppositions.

Automatisation des processus : Aussi appelée « automatisation opérationnelle » (et parfois même « transformation numérique »), l'automatisation des processus désigne l'utilisation de technologies numériques pour exécuter des processus opérationnels courants dans le cadre du déroulement du travail. L'automatisation des processus peut rationaliser une organisation pour la simplicité et l'amélioration de la productivité en libérant les humains des tâches simples et répétitives et en confiant celles-ci à des machines qui les exécutent plus rapidement. Un large éventail d'activités peut être automatisé, ou plus souvent, partiellement automatisé, où l'on maintient l'intervention humaine à des points stratégiques dans le déroulement du travail. Dans le domaine de la prise de décisions administratives à IRCC, l'« automatisation des processus » diffère du « soutien aux décisions automatisé », la première expression s'appliquant à des tâches administratives simples et la deuxième, à des activités où il faut faire preuve d'un certain degré de jugement.

Boîtes noires : Outils logiciels opaques fonctionnant en dehors de la portée de la surveillance et de la responsabilisation. Il s'agit souvent de systèmes d'apprentissage approfondi. Leur comportement peut être difficile à interpréter et à expliquer, soulevant des préoccupations quant au caractère explicable, à la transparence et au contrôle humain.

Code source : Programme informatique dans son langage de programmation original, lisible par l'humain, avant sa traduction en code objet. Il s'agit d'algorithmes et d'instructions informatiques et peut inclure des commentaires du concepteur.

Données massives : Terme familier désignant les énormes volumes d'information qui inondent les organisations au quotidien. Ces données sont souvent très variées et doivent être traitées au moyen de systèmes informatiques rapides; c'est pourquoi il est impossible de les gérer efficacement et d'en tirer de la valeur opérationnelle sans des outils logiciels et des méthodes d'analyse précis.

« Entrave » de l'exercice du jugement de l'humain : Il y a entrave lorsqu'un décideur n'exerce pas véritablement son jugement indépendant sur une question. Cela peut se produire lorsqu'un décideur se contraint à une règle fixe de politique, l'opinion d'une autre personne ou les résultats d'un système de soutien aux décisions. Bien que la personne qui prend les décisions administratives puisse être

favorablement influencée par des considérations politiques et d'autres facteurs, elle doit tenir compte des circonstances propres au dossier en question et ne pas s'accrocher aveuglément à un élément (p. ex. la note relative au risque fournie par un système algorithmique) en excluant d'autres facteurs pertinents.

Équité procédurale : Vitale pour la prise de décisions administratives, l'équité procédurale exige que les demandeurs :

- participent véritablement (ils connaissent le dossier et ont l'occasion de le défendre) et bénéficient d'une évaluation juste et exempte de préjugés de leur demande
- soient tenus informés des preuves négatives et des préoccupations des décideurs
- se voient offrir une possibilité réelle de répondre à de telles preuves et préoccupations à l'égard de leur demande
- se voient présenter les raisons de la décision

Il y a de nombreux facteurs qui déterminent ce qu'il faut aborder précisément pour respecter ces principes. L'un de ces facteurs est l'importance de la décision du point de vue du demandeur.

Gouvernance des données : Efforts déployés à l'échelle de l'organisation pour gérer officiellement les biens de données (par exemple, qui fait quoi, quand et selon quel pouvoir). Grâce à une gouvernance des données efficace, on s'assure que la qualité des données est supérieure au moment où elles sont créées, extraites et manipulées pour la production de rapports ou la prise de décisions. Faisant le pont entre le domaine des opérations et celui des TI, la gouvernance des données appuie l'innovation des processus et l'amélioration du service aux clients.

Intelligence artificielle : Englobant un large éventail de technologies et d'approches, l'intelligence artificielle est essentiellement le domaine des sciences informatiques consacré à la résolution de problèmes cognitifs associés à l'intelligence humaine, comme l'apprentissage, la résolution de problèmes et la reconnaissance de formes.

Prise de décisions administratives : Décisions prises par des responsables autorisés d'une institution (comme un conseil, un tribunal, une commission, un ministère ou un organisme du gouvernement ou un ministre) où les résultats touchent les droits, privilèges et intérêts. Élément central du système d'immigration, la prise de décisions administratives est également essentielle à d'autres domaines, comme le commerce, la fiscalité, la radiodiffusion et le transport.

Règles opérationnelles : Série d'instructions (propositions si-alors) qu'un système automatisé ne peut contourner. Des ensembles de règles peuvent être créés par des humains (selon l'expérience, le raisonnement, la recherche) ou maintenus par des humains (proposés par des modèles d'analytique avancée, mais examinés et approuvés par des humains). Les systèmes automatisés comprennent des règles opérationnelles qui ne peuvent être modifiées sans l'intervention humaine. Ils diffèrent des systèmes *autonomes*, qui sont assortis d'objectifs prédéterminés, mais d'aucune approche prédéterminée pour les atteindre.

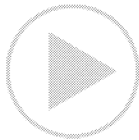
Système de soutien aux décisions automatisé : Inclut toute technologie de l'information conçue pour aider directement le décideur humain à prendre une décision administrative (par exemple, en formulant une recommandation) ou pour prendre la décision à la place d'un décideur humain. Il peut s'agir de systèmes comme l'automatisation des prolongations des permis d'études et des fiches de visiteur dans le SMGC ou l'autorisation de voyage électronique.

Zone d'exploration : La zone d'exploration, également appelée « bac à sable », désigne l'environnement servant aux fins de recherche, d'expérimentation et de mise à l'essai en ce qui concerne l'analytique avancée et l'IA. Les données, les codes et les logiciels sont isolés de ceux en production de manière à ce qu'ils puissent être mis à l'essai en toute sécurité

Immigration, Refugees and Citizenship Canada
Immigration, Réfugiés et Citoyenneté Canada

Policy Playbook on Automated Support for Decision-making

2021 edition



How to ensure responsible use of automation, advanced analytics and artificial intelligence in service of administrative decision-making

PREAMBLE

Why do we need new policy guidance in this domain?

The premise for automating aspects of administrative decision-making at IRCC is familiar: application volumes are rising in temporary resident, permanent resident and citizenship lines of business; operational networks are continuously required to expand their capacity; and clients are increasingly expecting quick and easy digital interaction. IRCC is undergoing a profound transformation, and new uses of technology are seen as one viable avenue to keep pace and still maintain quality.

Big data and artificial intelligence (AI) are exploding, and IRCC is beginning to recognize the potential they bring for new business insights and efficiency gains. But this is a new area, and conventional methodologies, guardrails and governance models are often incompatible with how algorithmic approaches really work. With algorithms and automation growing in importance, the Department would benefit from a considered approach to resolving novel questions about privacy, data governance, security, transparency, procedural fairness, and human-machine interaction.

The Government of Canada has introduced new government-wide policy around responsible and ethical use of automated decision systems, which came into force on April 1, 2020. The Treasury Board Secretariat (TBS) *Directive on Automated Decision-Making*, which supports the *Policy on Service and Digital*, outlines high-level requirements for departments using systems, tools or statistical models to recommend or make administrative decisions about clients. The Directive is be accompanied by a mandatory Algorithmic Impact Assessment tool, which must be initiated at the design phase and finalized before any new system moves from a test environment to live implementation. These overarching policy instruments will help IRCC to navigate uncharted waters.

To be successful, IRCC also needs clear and readily available guidance that takes into account its unique position as a service provider for both Canadians and foreign nationals, as well as its specific operating context, systems and programs. IRCC's internal policy starts with fundamental questions like: *Is the use of automation to support decision-making a good idea in my case? Do we have what we need to be successful? Do the benefits outweigh the risks? How do we do this securely?* It then grounds key technology ethics concepts in our day-to-day work, and directly answer the questions we are asking ourselves, such as: *Could we ever use algorithms to automate negative decisions on client applications? When, if ever, are "black box" algorithms appropriate? Who is accountable for individual decisions made by a machine? How do we meaningfully explain automated decisions to clients? How can we give IRCC's officers the benefit of insights gleaned from analytics without fettering their decision-making?*

IRCC needs to find the right answers to these questions to maintain public confidence, which is critical to successful immigration, refugee, citizenship, settlement and passport programs. Codifying or "hard-wiring" business rules, whether based on advanced analytics or on common sense heuristics long relied upon by individual officers, makes them subject to more scrutiny. The application of business rules at scale elevates the impact if something were to go wrong, requiring additional safeguards within the Department.

On top of managing anticipated risks, specific policy guidance on automated support for decision-making can help IRCC realize the full benefits of new approaches. Principles are intended to enable IRCC staff to consider not only what actions the Department should avoid because they are too risky, but also to what activities we should pursue because they can enable us to deliver *better* immigration programs and services.

What is the scope of activities this playbook targets?

This policy targets **automated systems that support, in whole or in part, administrative decisions.** This includes systems that: classify cases according to the level of scrutiny they require; flag cases for human review or investigation; provide recommendations about whether applications should be approved; or render complete decisions. The business rules applied by automated systems that support decision-making could be derived from sophisticated data analytics, or from interviews with experienced officers – what matters is that these systems take up a new role in IRCC's decision-making model. Analytics-based rules are a particular focus within this policy, especially if machine learning is involved.

This policy does not cover all types of automation. Routine process automation, such as the programming of computers to accomplish repetitive clerical tasks formerly done by humans, is out of scope. So too are systems that merely help to divide our collective caseload between networks, offices or officers for the purpose of efficiency. If the automation does not play a role in helping IRCC to reach an administrative decision that affects a client, it is not of primary concern in this policy. However, staff are encouraged to carefully consider whether automation that seems removed from final decisions may inadvertently contribute to an approval or a refusal.

What shape does the playbook take?

Full spectrum support for ethical conduct means having both a code of personal ethics and a set of institutional ethics. In terms of personal ethics, nothing new is needed. The existing Values and Ethics Code for the Public Sector is domain-agnostic (i.e. it already applies equally to accountants doing internal audits as it does to data scientists developing machine learning systems). The code's five core values of respect for democracy, respect for people, integrity, stewardship and excellence provide a sound and enduring basis for IRCC's innovators to make decisions about how to comport themselves ethically and professionally.

With respect to institutional ethics, a set of guiding principles is proposed in Part I. **As a general rule, principles do not identify specific courses of action, but instead serve as a basis for decision-making when people find themselves in new territory, confronting novel problems**. By disseminating the guiding principles in Part I, IRCC hopes to see its data scientists, program designers and policy developers inspired, trained and empowered to prioritize ethical considerations in the development of automated support for decision-making systems. In particular, they provide guidance to the Department as a whole and its senior managers, who are ultimately accountable for the results of automated systems and for creating a workplace that fosters responsible use of technology. Reflecting upon and openly discussing the wider impacts of the Department's innovative work can only help to ensure that it consistently supports the public good.

The principles presented in this policy are designed as a companion to the more concrete guide contained in Part II. **The Automator's Handbook guides IRCC staff through the primary questions that need to be considered at various stages of a project**: (1) when exploring automated support for decision-making as a possible solution to the problem at hand, (2) when designing a new system, (3) when preparing to launch, and (4) once an automated system is up and running. The handbook is geared toward individuals involved in developing and implementing automated systems that support decision-making. This may include data scientists, program and policy specialists, privacy experts, IT, and operational staff.

The Playbook covers additional chapters including **an overview of legal considerations and practical tips to align with administrative, human rights and privacy law (Part III)** and **baseline privacy requirements (Part IV).** The policy concludes with **a checklist to support compliance with the *Directive on Automated Decision-Making* (Part V)**, and a **glossary of key terms**.

Questions or comments about the Playbook? Contact IRCC's Digital Policy team in the Strategic Policy and Planning Branch: IRCC.Digital-Numerique.IRCC@cic.gc.ca.

A new policy on automated support for decision-making is an opportunity to ensure that the Department's thinking keeps pace with the speed of technological change, and that our people and practices continue to deliver a suite of programs equal to the expectations of Canadians and the world. The foundational principles and standards that have guided IRCC are not changing. Our commitments, mission and duties remain consistent. But while it may seem that the tools we use in the digital era merely recreate the paper and in-person practices of the past, they do not. Adoption of these technologies has and will change what we do and how we do it, the skills that we need, our relationship with partners and clients, and the culture of our workplace.

Note: This policy should be seen as a living document that needs to be regularly updated over time to ensure continued relevance as the technology, and our knowledge thereof, evolves.

PART I – A SET OF GUIDING PRINCIPLES

How should the Department as a whole proceed when faced with difficult new choices about specific applications of emerging technology? The following set of principles outlines IRCC's overarching goals and covers a range of topics, like the right tools for the right circumstances, responsible design, transparency and explainability, and recourse.

**Click the arrows on the left for an expanded treatment of each principle.*

Who needs to read this?

- *IRCC's senior managers;*
- *IRCC practitioners working in this domain; and*
- *Anyone interested in learning about the broad parameters that guide IRCC's use of AI and automation.*

1. The use of AI and automation should deliver a clear public benefit. IRCC should use these tools wherever it can do so responsibly, effectively and efficiently – in that order.

1.1 **IRCC's use of automated support for decision-making should build trust in Canada's immigration system**. Maintaining the confidence of Canadians is critical to the Department's success, and IRCC should actively address potential anxiety associated with Big Data and AI. The Department should take measured steps to build legitimacy and public trust around its use of automation, analytics and AI.

1.2 There is an opportunity cost to <u>not</u> pursuing the use of automation, advanced analytics and AI. IRCC should not be so cautious as to miss out on the potential of data-driven technologies. As in any era, **making use of the best tools available is part of effective and efficient use of public money.** In some cases, algorithmic systems are the only efficient way to process the massive amount of information needed for the operation of modern government services.

1.3 However, **IRCC can only employ these technologies to the extent that their use remains consistent with Canadian norms and standards** regarding equality and non-discrimination, procedural fairness, transparency, privacy, security, and accountability. The responsible use of technology must respect requirements of administrative law, and of overarching frameworks like the Charter of Rights and Freedoms and the *Privacy Act*.

1.4 **IRCC should not pursue automating decisions (or aspects of decisions) at the expense of program integrity.** Risk management is at the heart of IRCC's work. Automation may change risks or introduce new ones, but innovation should seek to boost program integrity.

1.5 **IRCC should make every effort to obtain reliable data** in order to make a direct link between clients' outcomes and the information they submit at the application stage. At the same time, **the Department needs to carefully consider the strength of the inferences it draws from new data sets**. Differentiating correlation and causation is a perennial challenge, and human intentions are hard to predict reliably.

1.6 When identifying potential areas to test the use of AI and automation, IRCC should consider the following factors:

- **Need:** Canada has seen year-over-year growth in the volumes of applications received across several lines of business, and expects this trend to continue for the foreseeable future. While growth is a primary driver for the adoption of AI and automation in IRCC's decision-making processes, enhanced program integrity management and rising client expectations are equally important. Focusing on need from the onset will ensure that IRCC is responding to a clearly identified problem and avoiding the introduction of new technologies for the sake of 'going digital';
- **Data:** the larger the traveller movement, the larger the corpus of cases to learn from, and where usage of online applications is widespread, data is more structured. The availability of large volumes of highly-structured data increases the quality of data, which can reduce unintended bias, and increase the integrity of data used to train algorithms and support decision-making. When exploring the use of machine learning to support temporary resident visa (TRV) processing, IRCC looked for caseloads where data integrity was high;
- **Stakes:** the gravity and finality of the decision, including its impact on clients, may vary by line of business. Temporary residence applications are generally considered lower stakes than permanent residence or citizenship applications, and certainly lower stakes than humanitarian and compassionate cases or pre-removal risk assessment; and
- **Complexity:** new technologies are more suitable for applications that involve routine examination of information, rather than cases that require more nuanced judgment calls.

See TBS Directive, section 4 (Objectives and expected results).

2. *Humans (not computer systems) are accountable for decision-making, even when decisions are carried out by automated systems.*

2.1 **The introduction of automated systems can change the *time and place* of human intervention in the decision-making process.** For example, humans may take on new roles, like setting business rules for an automated triage system to carry out. **However, automated systems should not displace the central role of human judgment in decision-making**. There is an essential role for humans in deciding which types of systems to use, which cases to apply them to, and which values to encode. In cases where algorithms play a role in administrative decisions that significantly affect people's lives, it is reasonable for applicants and Canadians to expect that a real person has exercised human judgment in developing the process, and that a person is ultimately accountable for its final results.

2.2 **Maintaining human intervention (often referred to as 'human-in-the-loop') is important for public confidence. Models that incorporate automation should maintain a role for people in the decision-making process.** Human-in-the-loop systems represent a form of transparency and accountability that is more familiar to the public than automated processes. This means that a

human decision-maker must be identified as the officer of record for all decisions made or supported by an automated decision system.

2.3 Accountabilities cannot be passed to machines when they take over tasks previously done by humans. **IRCC must take ownership for the successes and failures of its systems.**

2.4 IRCC should also reflect on the best approach to **devoting time to questions concerning the responsible use of AI** within its planning, advisory and decision-making structures.

See TBS Directive, sections 6.3.9-6.3.10 (Ensuring human intervention).

3. *Because IRCC's decisions have significant impacts on the lives of clients and Canadians, the Department should prioritize approaches that carry the least risk.*

3.1 Automating final decisions is the highest risk approach to achieving efficiencies. IRCC should first **explore whether efficiencies can be achieved by automating other aspects of the business flow** (e.g. workload distribution, risk-tiering, note-taking, research, communications, etc.)

3.2 When using a predictive model for decision-making, automating approvals is usually considered less impactful than automating refusals. However, it is important to view the situation from the perspective of individual applicants – particularly those applying to a business line where space is limited. **In situations where demand exceeds supply, the stakes for positive decisions are raised**: it may be perceived that an approval (or invitation or acceptance) of one candidate's application comes at the expense of another's.

3.3 Another perspective is that of Canadians, who accrue the economic, social and cultural benefits of temporary and permanent migration, but who also take on the safety and security risks. **From a collective perspective, automated approvals based on prediction can carry significant consequences**. It is understandable for Canadians to be more concerned about mistakenly approving risky applications than about mistakenly refusing *bona fide* candidates. IRCC needs to take a balanced view, considering fairness from both the individual and collective perspectives.

3.4 Because IRCC cannot fully eliminate risks to the security of its systems, it should **focus on building security measures into the design of its AI models and infrastructure to minimize impact in the event of a breach**. The mandatory Security Assessment and Authorization supports innovators in making sound decisions that enhance system integrity, management of sensitive information, and oversight.

See TBS Directive, sections 4 (Objectives and expected results); 6.1 (Algorithmic impact assessment), and 6.3.7 (Security).

4. *"Black box" algorithms can be useful, but cannot be the sole determinant of final decisions on client applications.*

4.1 **True "black box" AI tools are of limited utility for administrative decision-making** because the Department should not make decisions it cannot meaningfully explain. "Black box" tools, such as a

facial recognition application, can be used in a supporting role – but even then, **confidence/reliability levels should be sufficiently high; and where they are not, humans must have an opportunity to review the system's results before final decisions are made.**

4.2 **Black box algorithms should be permitted for the purposes of business optimization**, such as determining where to locate offices or how to best distribute resources among them.

See TBS Directive, section 6.2.3 (Providing explanations after decisions).

5. IRCC must recognize the limitations of data-driven technologies and take all reasonable steps to minimize unintended bias.

5.1 Every machine-learning algorithm operates entirely within the world defined by the data used to calibrate it, so **limitations or flaws in data sets will bias outcomes**, sometimes severely. In creating tools designed to make decisions on immigration, citizenship and passport applications, **IRCC must be careful not to replicate and "hard wire" historical biases** under a guise of technological neutrality. Understanding our data and planning for its responsible use is not a purely technical question – interdisciplinary collaboration is required to address fairness meaningfully.

5.2 AI-powered tools can also be used to study existing programs in new ways, to ensure they are achieving their intended results, and to **shine a light on any bias that may have inadvertently crept into long-standing practices** of human decision-makers.

5.3 All analytical tools are limited in their ability to predict and describe outcomes. While overall patterns and trends are useful, they cannot substitute for evidence of personal actions. For this reason, **IRCC should be very careful when applying information about a group to an individual**.

5.4 **Big data should be used to help IRCC look backward as well as forward.** IRCC should look to use new tools to detect, assess and remove any bias or inconsistencies.

5.5 **IRCC should constantly take steps to improve the quality of its data holdings, given that good data is foundational to advanced analytics and AI.** Prioritizing the development of data-driven tools requires prioritizing the breadth, depth and reliability of our data sets.

5.6 Analytics may suggest that an application is low-risk, based on how the application profile aligns with that of successful applicants in the past, and in this case **IRCC may wish to take on the risk of an automated approval. However, the calculus changes in the reverse scenario**: an automated refusal cannot be based solely on similarities between an applicant's profile and profiles of unsuccessful past applicants. **Refusals must be based on *evidence* about the particular application in question**. For example, an automated refusal based on a TRV applicant's passport being in Interpol's Stolen and Lost Travel Documents database would be justifiable because it is based on an objective, verifiable fact about the applicant in question (i.e. they do not meet the requirement for a valid travel document). Conversely, it would not be justifiable to base an automated refusal on the fact that the applicant's country of origin has a high incidence of passport fraud – this could serve as a risk indicator amongst others, but it is not itself a sufficient basis for an outright refusal.

See TBS Directive, sections 6.3.1-6.3.2 (Testing and monitoring) and 6.3.3 (Data quality).

6. Officers should be informed, not led to conclusions.

6.1 Automated systems that support decision-making can be used to assist officers in exercising their discretion. In these cases, systems must be designed so that they **do not fetter decision-makers** in the exercise of their capacities and authorities.

6.2 **Officers need appropriate training and context to understand what, precisely, a decision support system is telling them**. For example, labels matter: when a system automatically triages clients' applications into groups, it makes a difference whether these groups are labelled "green/yellow/red" or "straightforward/average/complex." The former may appear to represent instructions to officers (green=approve, red=refuse), the latter a simple indication of how much time and effort an officer will need invest to reach a decision.

6.3 Big data should **strengthen administrative decision-making by providing a richness of relevant information to decision-makers, and by focusing their attention on the determinative factors**. In this way, new tools should help officers make not only faster, but *better* decisions.

6.4 **IRCC should endeavour to provide its decision-makers with the best information and tools we have at our disposal**. Putting new insights at the fingertips of decision-makers and focusing their attention on the most relevant factors will reduce wasted effort, boost consistency and better link action and rationale.

6.5 User-testing during the design and development of a project can enable the Department to **develop a strong grasp on how humans interact with automated systems**. Beyond providing insights into the most effective and reliable ways to present machine-generated outputs to officers, a more in-depth understanding of the dynamics at play can enhance accountability and help prevent unintended negative consequences.

See TBS Directive, section 6.3.5 (Employee training).

7. Humans and AI play complementary roles. IRCC should strive to sharpen the roles of each.

7.1 Automation, advanced analytics and AI are force multipliers; they can **assist and augment the capabilities of people**. Using these technologies to allow IRCC staff to focus on more creative, problem-solving, or client-facing work should improve both service delivery and job satisfaction.

7.2 **IRCC should leverage the strengths of computers** – reliably analyzing large volumes of data and considering millions of possibilities without fatigue – **and let humans focus on theirs**: intuition, creativity, empathy, social skills, shaping a larger strategy. **Algorithmic systems should allow people to focus on things that require their expertise and judgment.**

7.3 **Systems that make administrative decisions on their own** – as opposed to assisting a human decision-maker – **are suitable for decisions involving straightforward, factual determinations**, such as whether a client has a criminal record or has visited Canada before. Automation should be

focused first and foremost on routine tasks – where desired outcomes are easily described and where human judgment is not required.

7.4 In contrast, when automated systems are intended to assist a human decision-maker, training should be designed in a way that encourages officers to **apply their own professional judgment and intuition**, as a means of augmenting decision-making[1].

7.5 **Having a human involved in the process not only safeguards against procedural fairness risks, it gets better results.** For straightforward types of assessments (e.g. photo matching), the performance of data-driven algorithms is often superior to that of humans. But when machines and humans team up, the results are better still. One way to boost performance is to have experts apply judgment to the output of mathematical models; another is to flip the sequence and have program experts and decision-makers contribute to the design of the model in the first place.

7.6 **IRCC should regularly review and adjust its automated systems** to ensure they reflect the reality on the ground and are operating in alignment with the policy priorities of the day. Feedback loops are critical to responsible design and quality control. For example, feedback from frontline officers must be able to reach system developers, so that IRCC can constantly reassess whether a system's conclusions are reliable. **It is essential that officers continue to review a portion of applications manually**. This will help to ensure that officer skills do not atrophy and that knowledge about country conditions and fraud trends remains up-to-date.

7.7 **Technical experts (and their managers) are in a position of trust, and have a special responsibility to thoroughly evaluate the business logic and the potential issues that may arise due to automation.** It is imperative that the Department receive objective assessments and frank advice. With the advent of machine learning systems, much care should be taken to identify, explain and mitigate potential risks, especially when risks may only be apparent to those with a technical understanding of those systems. Managers need to create conditions that allow experts to consult peers, reassess risks as systems evolve and report hazards without fear of reprisal.

7.8 IRCC should make every effort to **give its employees the skills, tools and support they need** to pursue innovation. Support includes time to learn, and the use of secure sandboxes or testing environments that allow employees to experiment responsibly.

See TBS Directive, sections 6.3.9-6.3.10 (Ensuring human intervention) and 6.3.5 (Employee training).

8. IRCC should continually adopt emerging privacy-related best practices in a rapidly evolving field.

8.1 Data scientists, program designers and IT specialists should be conversant in privacy issues, including **understanding the rights and responsibilities associated with the collection, use, disclosure and retention of personal information**, as well as related **best practices and standards** IRCC should give

[1] https://www.nesta.org.uk/blog/human-vs-machine-why-we-need-be-thinking-more-about-how-humans-are-interacting-ai-tools/

thought to developing tools and procedures to assist staff in assessing and addressing privacy needs in the age of big data.

8.2 The Department should **be proactive in preparing for a possible privacy breach or investigation into how it manages personal information** in its algorithmic systems. This means following appropriate security and privacy guidelines regarding breaches, and ensuring that training is current. In the event of a complaint or review by the Office of the Privacy Commissioner, IRCC must be prepared to produce copies of all documentation outlining the steps taken to comply with privacy requirements, and steps taken to exceed mere compliance, where appropriate.

See TBS Directive, section 6.3.3 (Data quality).

9. IRCC should subject systems to ongoing oversight, to ensure they are technically sound, consistent with legal and policy authorities, fair, and functioning as intended.

9.1 Inviting **external scrutiny** of the design of our systems is an important aspect of ensuring legitimacy. Where possible, new algorithmic models should be made available to external experts, such as an advisory body, for auditing, testing and review.

9.2 **Quality control** processes should be established early on and conducted regularly. If testing is done internally, the methods and assumptions used, along with the results, should be openly documented and made available to the advisory body. Subjecting automated decision systems to greater control mechanisms immediately prior to deployment and on an ongoing basis post-deployment while providing a lighter touch during exploration will ensure adequate oversight without stifling innovation.

See TBS Directive, sections 6.3.2 (Testing and monitoring) and 6.3.4 (Peer review).

10. IRCC must always be able to provide a meaningful explanation of decisions made on client applications.

10.1 Regardless of the degree to which IRCC relies on algorithmic models to support decision-making, applicants will always have the right to understand the basis of decisions made on their applications. Explanations play a key role in allowing applicants to pursue recourse options, and to government accountability and public trust more generally. Explaining how algorithmic systems work can be technically challenging, but **explanations of automated decisions should not hinge on the general public's understanding of how algorithmic systems function**. If an explanation provides little meaningful information, it offers little value.

10.2 One way to make explanations understandable is to limit a system's performance to those functions that can be easily explained. This may, however, mean squandering the potential of AI systems that can make sense of complex data in ways that humans cannot. **Rather than imposing sharp limits on what new tools can *do*, IRCC should endeavour to find new and better ways to *explain* them.**

10.3 Given a prohibition on black box systems for administrative decision-making in full, the use of automation to support decision-making should make decisions **more explainable, not less**. For example, audit trails capturing every factor and step in the decision-making process can be used, where needed, to systematically generate and record reasons for acceptance or refusal.

10.4 An **audit trail** is critical to respecting an applicant's right to understand the basis of the decision on their application, and may be needed to satisfy the courts in the context of a legal challenge. Any new tool that lacks the capacity to **automatically generate a comprehensive audit trail of the administrative decision-making path** is not yet ready for implementation.

See TBS Directive, section 6.2.3 (Providing explanations after decisions).

11. IRCC must be transparent about its use of AI. It must provide meaningful access to the system while protecting the safety and security of Canadians.

11.1 **IRCC needs to find ways to report, in a meaningful way, on the workings of whole systems.** At a minimum:

- Clients need to understand how their applications will be processed and their personal information used;
- Canadians need to understand how IRCC spends public money to facilitate the entry of family, friends and legitimate visitors, while at the same time protecting against threats to health, safety and security; and
- Stakeholders, media and technical experts need a basis for an informed critique of Canada's approach. This means that IRCC should proactively share information about how its models work and, more importantly, how the Department is – and is not – using new technologies to support decision-making.

11.2 The Department has an interest in not disclosing too much information about its decision systems for program integrity reasons. Divulging too much information can allow individuals to manipulate decision-making, or "game the system." Explanations of how systems work, and of the individual decisions they help the Department to make, should be **detailed enough to inform a client generally about how decision will be made, but not so comprehensive as to jeopardize program integrity and, by extension, the safety and security of Canadians.**

11.3 **IRCC should endeavour to design systems, to the greatest extent possible, around factors/data points that are difficult to falsify or misrepresent.** For example, bank records are much more difficult to fabricate than a reason for travel, which is simply stated by the applicant. We should rely on objective facts as much as possible in order to mitigate potential fraud risk. The weight assigned to different data elements in IRCC's algorithms should account for the degree to which they may have been targeted for fraud in the past, or are likely to be targeted going forward. For example, potential security risk factors would be highly weighted.

See TBS Directive, sections 6.2.1-6.2.2 (Providing notice before decisions) and 6.2.6 (Release of source code).

12. IRCC's use of automated systems must not diminish a person's ability to pursue recourse.

12.1 When a client challenges a decision that was made with the assistance of an automated system, and IRCC has the authority and agrees to review the decision, **a human officer will conduct the review**. Similarly, under the Privacy Act, individuals have a right to recourse – via a complaint to the Office of the Privacy Commissioner – if they feel their personal information was mishandled.

12.2 However, **IRCC should not proactively offer clients the choice to have a human officer review and decide on their case in lieu of the automated system at the beginning of the application process**. IRCC's objective is to implement responsible, effective and efficient systems. Allowing clients to opt out of them would compromise the Department's ability to ensure consistent and fast processing, while also erroneously implying that its automated systems provide inferior service.

See TBS Directive, section 6.4 (Recourse).

PART II – THE AUTOMATOR'S HANDBOOK

This practical handbook is intended to guide IRCC staff through key questions that should be considered at various stages of AI and automation projects – from early exploration to ongoing monitoring once a system is up and running.

Records of key decisions about the use of AI and automation is important to demonstrating that you have carefully considered a host of factors and identified adequate mitigation strategies along the way. Keeping notes of observations, analysis and links to key documents throughout the design, development and deployment phases can be helpful to support the answers you will need to provide when you complete an Algorithmic Impact Assessment (AIA), and may also need to be produced in the event of litigation. More information about the AIA tool and requirements is provided in section B, *Designing your system*.

Who needs to read this?

- *Individuals and teams who are considering the application of AI and automation to support administrative decision-making. This may include data scientists, policy and program leads, privacy and data governance professionals, IT specialists, and operational guidance experts – most likely at the Analyst, Officer, Advisor, and Manager levels.*

A. Exploring automated support for decision-making as a potential solution

This section will help to work through the key questions that need to be considered at the outset of a project. The answers to these questions should help you determine whether or not an automated decision system is a suitable solution to the problem you are trying to tackle, whether or not you have what you need to be successful, and if the benefits outweigh the risks.

General suitability

Have you clearly defined the problem at hand?

Spending time focusing on the problem before jumping to possible solutions will help avoid 'technology theatre' and make sure that the solution you ultimately choose responds to a clear need. It can also save time and help you identify relevant metrics to measure success.

Do all stakeholders have the same understanding of the problem? Is it an issue that can be solved through AI and automation?

Add observations, analysis and links to key documents here.

Have you exhausted straightforward options before moving into the more complex and contentious world of automated support for decision-making?

In some cases, automation of simple and repetitive tasks (e.g. through the use of robotic process

automation) or new approaches to workload distribution may be able to achieve similar efficiencies. Considering multiple options upfront.

Add observations, analysis and links to key documents here.

Does the nature of the decision or specific step in the decision-making process lend itself to automated decision-making?

When determining whether a task or step in a decision-making process could be shifted from humans to machines, you should ask yourself, "Is this a situation in which reasonable minds may differ?" If yes, automation is ill-advised. Conversely, if analysis of past decisions has shown that virtually any officer would reach the same conclusion given the facts at hand, automation should be pursued.

Add observations, analysis and links to key documents here.

Is this a project that requires special attention? For example:

- It is within an area of intense public scrutiny (e.g. because of privacy concerns) and/or frequent litigation
- Clients in this line of business are particularly vulnerable
- The stakes of the decisions are very high, e.g. you are thinking of automating refusals, or outcomes will have a significant impact on peoples' lives
- It will have a major impact on IRCC staff, either in terms of their numbers or their roles
- You are thinking of collecting/using non-traditional and potentially sensitive data, such as data associated with clients' social media accounts

If one or more of these points applies, your project will require more consultation and oversight at all stages.

Add observations, analysis and links to key documents here.

Is the technology you are thinking of using likely to raise public concern? For example:

- The specific type of technology is a source of negative media coverage or associated with outcomes that are less effective when it comes to certain minority groups (e.g. facial recognition)
- You are thinking of using a "black box" algorithm
  - *How do you know if you have a "black box"? The system's process of analysis or reasoning is unknowable, or its behaviour is difficult to interrupt or explain*

If so, you will need to consider how you will ensure that the introduction of these tools does not negatively impact confidence in IRCC's ability to make decisions that are fair and reliable.

Add observations, analysis and links to key documents here.

Preliminary diagnostics and impact assessments

What impacts, positive and negative, will the proposed automation have on clients?

For example, the project is expected to:
- Lead to faster processing
- Require clients to entrust IRCC with more personal information, or otherwise affect their privacy
- Improve or degrade the quality of information provided to clients with respect to decisions rendered
- Result in some cases or applications receiving little or no scrutiny by a human decision-maker

Add observations, analysis and links to key documents here.

Will your automated decision system impact diverse groups?

Gender-Based Analysis Plus (GBA+) is a comprehensive analytical tool that helps fully assess the potential impacts of new government policies and initiatives on diverse groups of people. The "plus" acknowledges multiple, intersecting factors such as: sex and gender, indigenous status, income, employment status, family status, education, immigration status and country of birth, age, visible minority status, language, region of residence, disability status.

Applying a GBA+ lens at the very start of your project will help you assess risks that may result in differential impacts for diverse groups. At this stage, you should focus on gathering disaggregated data about clients, analyzing this data for quality and historical bias, and checking your assumptions. AI can also be useful to establish relevant patterns of intersectionality, which may go unnoticed by humans. GBA+ analysis will also come in to play at later stages of your project!

This is the time to engage IRCC's GBA+ Unit, located in the Strategic Policy and Planning Branch, as early as possible, for advice and available training.

Add observations, analysis and links to key documents here.

What impacts, positive and negative, will the proposed automation have on the Department?

For example:

- Reducing existing backlog of cases
- Saving money (in the short, medium or long term)
- Improving consistency and overall quality of decisions
- Improving program integrity
- Performing tasks that humans could not accomplish in a reasonable amount of time, allowing officers to focus on more complex tasks requiring human judgement
- Increasing the risk of litigation or privacy breaches
- Increasing (or decreasing) the robustness of data being collected and used

- Changing the role of officers and decision-makers, including their core competencies, job descriptions or classifications, or changing the overall size or make-up of the staff required to deliver a program

Add observations, analysis and links to key documents here.

Have you considered how frontline staff might react?

You may wish to consult with them, and with Human Resources. Even proposals that do not entail substantive changes to officer roles can inadvertently provoke uncertainty and trepidation. Efforts to communicate project goals up-front could reassure affected staff, saving time and effort later.

Consider how you will consult with frontline staff to get their feedback on a proposed automation tool. This would provide an opportunity to convey the objectives of the proposed automation tool and could help ensure concerns can be addressed as early as possible. The section on adopting a User-Centered Approach offers more detailed guidance.

Add observations, analysis and links to key documents here.

Have you identified potential legal implications and risks?

In the initial stages of your project, you will need to establish a diagnostic of legal issues and flag potential legal implications for further review. This should include all areas of public law (i.e. administrative law, human rights law and privacy law). See the guidance below on *Legal Considerations and Tips* for more information on how to assess and mitigate legal risk throughout the AI lifecycle.

IRCC's Legal Services Unit has a team of counsel specialized in questions related to the use of technologies like AI and automation. You must contact Legal Services early in the process and on an ongoing basis to seek advice pertaining to your project.

Add observations, analysis and links to key documents here.

Will you require new policy or legal authorities or regulations?

The Department has broad legal authority to use electronic tools to deliver on its mandate related to immigration and passport decision-making (see section 4.1 of IRPA and amendments to section 2.2 of the Canadian Passport Order), but you could still require new authorities if your project would entail automation in other lines of business or substantive changes to the following: program requirements (e.g. eligibility criteria); program outcomes (e.g. which clients are approved); the collection, use or disclosure of personal information; or the role of officers in the decision-making process. Even if you have the legal authorities to introduce the proposed automation, regulations may be needed to provide guardrails or to specify exactly how and when automation will be used.

You may need to work with legal and policy leads to seek new authorities or to introduce regulations through a Memorandum to Cabinet or other means.

Add observations, analysis and links to key documents here.

Does IRCC have the breadth and quality of data required to make the proposal workable?

Establishing a diagnostic related to data management will not only help you identify and mitigate legal risk, it will also help you assess the feasibility of your project. In doing so, you should consider data management as it relates to collection, use, disclose, retention, and security of the data, including personal information (which may require a privacy assessment). You should also consider whether you will need to develop a data strategy for the project.

Some of the questions to consider are:

- Is there enough high-quality, structured data to support reliable inferences? What are the limitations of the data? Is it accurate, up-to-date and truly representative, and does it contain historical biases that could be perpetuated? You will need to get an assessment of the quality of your data set, with leadership from the departmental data governance function and collaboration between business and IT. This assessment should be approved by a data governance management table, such as the Data Executive Steering Committee. *CDO Branch to confirm if this is still valid/up-to-date*.
- How many of your data points are proxies for the actual criteria or outcomes you are targeting? All of them? You will need tried and tested indicators that applicants comply with program objectives/requirements.
- Can you use the data? Does the model make inferences about personal information not provided by the applicant? Does it invade or pose significant risks to individuals' right to privacy? On what basis would you use the data underpinning your analytics in a real world application? Do you only need access to information that IRCC already collects (cross program uses of information previously collected for one program should be reviewed), or are you seeking to pull in external data, such as that held by federal, provincial or international partners?
- How will the data be collected and stored, in relation to other IRCC data systems and data assets?

You will need to get advice from the CDO Branch, the ATIP Division and possibly Legal Services about whether your proposed collection, use or disclosure of the data is consistent with the privacy framework governing the Department, or whether new authorities should be sought. The GBA+ Unit can also provide advice around the collection and display of sex or gender information.

Add observations, analysis and links to key documents here.

Have you assessed the privacy needs related to your project?

Generally speaking, you will need to consider what personal information may be used within your project and how it could be used with the lowest privacy interest. You will also need to evaluate data flows to ensure collection, use and disclosures of personal information meets legal and policy

requirements. Familiarizing yourself with 'Privacy by Design' principles will help you think about the possible privacy considerations and best practices that could help make your proposal robust.

IRCC's Integrated Corporate Business Branch has developed new internal tools and procedures to enable IRCC staff to adopt a "privacy sensitive" approach to projects, including those that involve AI, advanced analytics and automation.

Completing a **Privacy Needs Assessment** (internal link) is the first step to determine what measures you will need to take to adequately mitigate and manage privacy risks related to your project. Questions focus on legal authorities for collection, specific types of personal information that will be collected or created, who will store and have access to this personal information, and how the data will be used (e.g. cloud computing, chatbots, analytics, AI, web scraping, data matching). Once you have completed the form, an ATIP advisor will review it and may recommend:

- The development of a new Privacy Impact Assessment or PIA (internal link);
- An amendment or update to one or more existing PIAs; or
- Other privacy work, as appropriate.

Add observations, analysis and links to key documents here.

Training

Have you taken advantage of available training related to data-driven technologies and their effects?

This may range from courses on digital government – such as those offered by the Canada School of Public Service's Digital Academy – to training on privacy and data literacy. Training on GBA+ will also help you unpack questions around bias and identify actions to avoid negative unintended consequences on diverse groups.

While not all members of a team working on an automated decision system will need the same level of expertise in all areas, the TBS *Directive on Automated Decision-Making* requires that adequate training be provided to employees in the design, function, and implementation stages to ensure that they are able to review, explain and oversee the operations of an automated decision system.

Add observations, analysis and links to key documents here.

Partner and stakeholder engagement

Have you thought about which internal subject matter experts needs to be engaged in the project?

Ensuring the responsible use of automation in decision-making requires vast expertise in numerous areas. Bringing together a multi-disciplinary group will help you tackle the policy, the legal, the data, the IT, and the operational questions all once. It will also help you identify relevant strategies to mitigate risk (e.g. bias in your automated system), create greater awareness about your project throughout the Department, and facilitate alignment with other related initiatives.

You will likely need to involve subject matter experts from the following teams:

- **Advanced Analytics Solutions Centre (A²SC), Operational Planning and Performance Branch** – in most cases, A²SC's team of in-house experts in applied data science will be responsible for designing and testing automated models that support decision-making.
- **Legal Services** – tech counsel can provide advice on program requirements, privacy, Charter, administrative law, intellectual property and litigation risks. Under the TBS Directive, you are required to consult with legal services from the concept stage of a project to ensure that the use of the automated decision system is compliant with applicable legal requirements.
- **ATIP Division, Integrated Corporate Business Branch** – the Privacy, Policy and Governance Unit can provide advice on privacy considerations and requirements, including how to complete a Privacy Needs Assessment, whether you need to undertake or revise an existing Privacy Impact Assessment (PIA), privacy notice statement or Personal Information Bank (PIB). The team can also work with you to build privacy into individual projects.
- **Digital Policy team, Strategic Policy and Planning Branch** – this team can provide advice on policy and legal authorities, IRCC's and TBS' policies on the responsible use of AI and automation, connections with other departmental initiatives, etc.
- **GBA+ team, Strategic Policy and Planning Branch**
- **A program policy branch** – depending on the use case, you may need to work with program policy areas to ensure alignment between your initiative and future policy directions, and seek new authorities (e.g. Immigration Branch, Citizenship Branch, Settlement and Integration Policy Branch).
- **Chief Data Office Branch** – as the departmental lead for 'data quality by design,' this branch can help assess data quality and its suitability for your project, and provide advice on questions related to data governance (e.g. data quality, data management, data lifecycle, and business glossaries). This team should be involved in the design, development and implementation of automated systems.
- **Immigration Program Guidance Branch, Integrity Risk Management Branch, and processing networks** – these teams can provide first-hand knowledge of decision-making processes, the role of officers, related guidance, and operational impacts related to your project.
- **Digital Strategy Branch** – this team can help assess how your initiative fits within IRCC's digital IT strategy and business capabilities.
- **Communications Branch** – this team can support you in determining whether there is a need to do any public communications regarding your project.
- **Client Experience Branch** – this team can support the development of end-to-end human-centered design and usability testing related to your project.
- **Human Resources Branch** – this team can assess whether a new automated system will require modifications to generic work descriptions, statements of merit criteria or position classification requirements. HRB can also provide advice on when and how to engage unions.
- **IT Operations, Transformation and Digital Solutions Sector (TDSS)** – this branch is responsible for system design and architecture, and provides advice to meet the IT security needs of the Department, ensure compliance to applicable laws, policies and standards. You will need to engage the IT Security Unit during the design phase to initiate a Security Assessment and Authorization process, which is mandatory under the Policy on Government Security.

Given the scope and complexity that may accompany automation proposals, you will likely also want to think about how you will engage your management team and other senior officials throughout the development of your initiative. This will help ensure senior management has all the information they need to make decisions at the right times throughout the process.

Add observations, analysis and links to key documents here.

Have you thought about whether you need involvement from other government departments (e.g. CBSA) or provincial/territorial governments? Although this policy targets IRCC activities, your project could have impacts on partners (who may also have their own policies/approaches regarding analytics, AI and automation).

Add observations, analysis and links to key documents here.

Will you engage external stakeholders?

External engagement around IRCC's use of AI and automation in decision-making is vital to building public trust, ensuring that there is not a vast disconnect between how IRCC uses these technologies and what the public views as permissible, and, ultimately, enhancing our capacity to advance responsibly in this domain.

It may not be practical to consult external stakeholders on the development of every algorithm or automated system used by IRCC, but they can be consulted on the potential impacts. It is important to consider the perspectives of those who will be affected by automating aspects of the decision-making process, and particularly when new algorithms/systems are part of an entirely new program or a substantial change in delivery approach.

Traditional IRCC stakeholders are still important, but you will likely need to consider a broader array of stakeholders, such as those working in the field of AI/technology ethics.

It is recommended that you deliberately seek views from a diverse group of stakeholders, and document their perspectives as you would when developing a significant policy or legislative change.

Add observations, analysis and links to key documents here.

Planning for design

Have you considered what resources will need to be budgeted for data analytics experimentation and iterative systems development in your Memorandum to Cabinet and/or Treasury Board Submission (if required)?

Uncovering new business insights through the use of advanced analytics takes time and costs money. Building a new decision support system can be a process of trial and error, and budgeting should provide the flexibility to iterate as needed before achieving a final product. Once built, automated systems require resources for maintenance and periodic renewal.

In the early stages of the project, you will likely need to present the proposed approach to the Business Intake Board (BIB). Led by Projects Branch, this centralized intake process for IM/IT investment proposals (and some non IM/IT projects) assesses the business value and prioritization of new ideas

from a departmental perspective. It also helps ensure that new investments support the strategic direction set out by IRCC's priorities and transformation roadmap.

Add observations, analysis and links to key documents here.

Have you considered how resource requirements will change over time?

It is common to need more human and financial resources at the outset, when a rigorous quality assurance regime is vital to ensuring the system is functioning as intended. It is also important to forecast ongoing costs related to system maintenance, monitoring, and renewal.

Add observations, analysis and links to key documents here.

Will your automated decision system be designed in-house or by an external supplier?

In most cases, the design and development of models will likely be done in-house, by A^{2}SC. In the event that projects are out-sourced to third parties, you should consult A^{2}SC for advice and technical expertise. You should also review Government of Canada guidance on using third-party vendors when it comes to AI and automation (e.g. Guideline on Service and Digital)

Public Services and Procurement Canada and TBS have established a list of pre-qualified suppliers interested in providing AI solutions to the Government of Canada. You will also need to carefully consider TBS requirements related to access and propriety of source code.

Add observations, analysis and links to key documents here.

User-centered approach

While human accountability and human-in-the-loop models are familiar concepts in the domain of automated decision-making, the idea of an end-to-end user-centered approach is less so.

And yet, failure to consider the needs and realities of users can increase risks ranging from legal liability (as a result of undue reliance or fettering discretion) to diminished returns on investment (because new systems are not well integrated into existing operating practices), and lower-quality decisions (because users lack an adequate understanding of the system and its outputs).

Have you taken steps to understand users' needs and their current environment?

Prior to undertaking a project that could have a substantial impact on the role of officers in decision-making, you will need to have a good understanding of the general operating environment in question.

- Establish who the primary users are and engage them directly to understand their needs. The users of IRCC's automated systems will generally be officers in our operational networks. Use a combination of open calls for comments and targeted solicitation to gather diverse views.
- Identify existing pressures, pain points and limitations, and factors that might support or impede change. This may include constraints related to time, resources or IT infrastructure.

- Anticipate impacts (both perceived and real) of automation on users, recognizing that all users are unlikely to be affected in the same way.
- Familiarize yourself with the tools and processes that are already in use (or planned) throughout the decision-making process and in targeted offices and/or lines of business. This will help ensure that you can develop compatibility and align your system in a way that makes sense to users.
- Consider your own beliefs and potential biases about the current operating environment, and how they may affect your interaction with users. This should be included as part of your broader GBA+ analysis.

Add observations, analysis and links to key documents here.

B. Designing your system

This section will help to answer questions like: *How do we design for fairness? How do we safeguard the impartiality of our decision-makers? How do we meet client expectations for transparency and protection of personal information?*

You will also need to provide some of the answers when you complete the Algorithmic Impact Assessment (AIA) – more information about the AIA is provided below.

Model suitability

Have you modeled more than one option?

It is unlikely that there is only one way to achieve your goal. Exploring, mocking up and testing some alternatives might confirm your initial hypothesis and strengthen your business case – or it might help you to uncover an even better approach.

Remember that the privacy framework applies even in the modeling stage of your activities. The ATIP Division can provide information on specific requirements, and advice to help you adopt a privacy-sensitive approach.

Add observations, analysis and links to key documents here.

Have you developed a detailed process map that will enable policy specialists, privacy experts, legal counsel and others outside of your domain to understand the business process in question, and how, specifically, the introduction of automation will change it?

Add observations, analysis and links to key documents here.

Algorithmic Impact Assessment (AIA)

Have you completed a preliminary AIA?

The AIA is a questionnaire designed to help you assess and mitigate risks associated with deploying an automated decision system. It includes roughly 60 questions about the project, algorithm and data, as well as de-risking and mitigation measures for data quality, procedural fairness and privacy.

Completing a preliminary assessment at the design stage will help you anticipate risks associated with your project, identify mitigation measures and plan to comply with requirements related to:

- Peer review
- Notice
- Human-in-the-loop for decisions
- Explanation requirements
- Testing
- Monitoring
- Training
- Contingency planning

– Approval for the system to operate

Determining the Department's responses to the questions in the AIA will require the perspectives of multiple internal subject matter experts. As such, you should engage policy and program representatives, legal counsel, data governance and data scientists to complete this assessment, and keep records of any supporting documentation or analysis. You may also need to obtain approval from certain branches (e.g. CDO Branch signoff on questions about input data and data quality).

You will need to update the AIA to reflect any substantial changes to your model, and release the final results publicly.

Add observations, analysis and links to key documents here.

User-centered design

Have you engaged officers in designing your system's user interface and thinking about the most effective approach for human involvement?

Getting the most out of automation tools requires thinking about both the most appropriate way to use technology and about the best way to involve human. Focusing specifically on human-machine interaction will help strengthen accountability, safeguard the impartiality of decision-makers, and strike the right balance between the roles of humans and the system. Engaging users directly in the design of these processes (rather than simply informing them after the fact) will build trust and confidence in the automated system. Effective co-design should also draw on rigorous experimentation and an iterative approach – i.e. start small, test different approaches, gather insights, adapt the design, and repeat.

- Experiment with different ways of presenting machine information to officers, using a rigorous approach to measure the impact of various designs. For example, using colour-coded labels to triage applications may prompt different types of responses than presenting information labelled with a percentage or level of complexity. Poorly designed machine outputs and human-machine interaction can mask issues of automation bias and fettering decision-making.
- Give careful consideration to ensuring that the system's user interface and outputs encourage users to exercise their own judgment.
- Work with officers and other multidisciplinary team members to map out the key touch points and tasks requiring human intervention. You should also consider the circumstances in which an officer would be authorized to override a decision made by the system.
- Design and deliver training that gives officers a full understanding of the capabilities and limitations of the system.

Add observations, analysis and links to key documents here.

Fairness and non-discrimination

What steps will you take to minimize unintended bias?

Poorly designed algorithms can introduce bias, but in most cases bias comes from the data that is used for analytics or for the training of machine learning models. Program designers need to be aware of weaknesses in the data at the root of their project. Some techniques for mitigating bias in data include consulting subject matter experts to ensure the data is situated in the appropriate context (e.g. has IRCC changed any business definitions over time?), and subjecting machine learning models to adversarial training that could correct for perceived unfairness. You may also want to consider having external partners or experts review your proposed data sets to help identify any gaps or possible concerns. Rigorous quality assurance testing of results can also detect trends that suggest bias, allowing for immediate correction.

Whatever the specifics of a given project, it is always advisable to assemble a diversity of views and identities within your development team. There are numerous examples of AI projects gone wrong due to a lack of diversity within the brain trust. Overlooking the unique circumstances faced by different groups of people can have damaging impacts on the populations your system is designed to assist. Guarding against unwanted bias means bringing a GBA+ lens to your data analysis and early design work. You should endeavour, to the greatest extent possible, to have team members who bring different worldviews, be they based on sex, gender, age, ethnicity, language, rurality, ability or other identity factors.

Add observations, analysis and links to key documents here.

How will your system be designed to ensure procedural fairness?

In order to preserve (or enhance) procedural fairness, you will need to think carefully about how the addition of automation will change application processing and decision-making. For instance:

- If you are planning a predictive model, consider how it will inform officers without fettering their discretion.
- Ensure that all relevant factors are still being considered.
- If your system will draw on information beyond what an applicant submits, consider how clients will be informed. Clients have a right to know what other ("extrinsic") evidence was used in decision-making, in part so they have the opportunity to correct any erroneous information.
- Establish the reasons that will be provided when a decision is made.

Add observations, analysis and links to key documents here.

Explainability and transparency

How will you ensure that meaningful explanations are provided to clients?

The Government of Canada's *Directive on Automated Decision-Making* requires that meaningful explanations be provided to affected individuals of how and why the decision was made. Specific parameters vary depending on the level of impact associated with the automated system.

As a general rule, explanations of automated decisions should: (1) help clients understand why a particular decision was reached, and (2) provide grounds to contest the decision should the client wish to do so. Program designers should also consider whether it is possible to indicate, at a high level, what would need to change in order for the client to achieve a desired result in the future. This is sometimes called a "counterfactual explanation."

IRCC has yet to pursue automated refusals of client applications. If the Department were to do so, we would need to be mindful of the following:

- Administrative decisions made by, or with the assistance of, an automated system must be at least as explainable as the decisions traditionally made by humans alone. This is why "black box" algorithms, like advanced neural nets, are not recommended as a primary way of determining whether to approve an application.
- Explanations must not jeopardize program integrity. A level of transparency that would include disclosing the inner workings of an AI system is not generally necessary. Instead, explanations are mainly required to answer questions like: What was the basis of a decision? What were the main factors? Would changing a certain factor have changed the decision? Why did two similar-looking cases lead to different decisions?

Add observations, analysis and links to key documents here.

Considering the need for transparency, particularly in the context of legal challenges, are you proposing business rules that are defensible and difficult for individuals to "game"?

- Models and rules need to address program requirements such that decisions are based on these requirements and on relevant, reliable information.
- Design your system, to the greatest extent possible, around factors/data points that are difficult to falsify or misrepresent. Much of the information on client applications is simply stated by applicants, with no way for IRCC to authenticate, but other data points, such as bank records or previously issued visas, are verifiable. Preference for the latter will help mitigate the risk of "gaming by claiming" on the part of applicants who start out with knowledge of how IRCC's systems work.
- If weighting variables in an algorithm, your weighting should account for the degree to which each variable may have been targeted for fraud in the past, or is likely to be targeted in the future.

Add observations, analysis and links to key documents here.

Privacy

How will your system uphold the privacy rights of clients and Canadians?

Consider Privacy by Design principles throughout development to help ensure privacy is built in from the start.

Your approach to using personal information for legitimate ends, without violating the rights of individuals or groups, should include:

- Understanding the provenance of the data
- Limiting the collection of personal information to only what is necessary
- Ensuring its accuracy
- Reducing unnecessary, redundant or marginal data
- Taking precautions to prevent re-identification of anonymized (or pseudo-anonymized) data or unauthorized data collection
- Protecting data from unauthorized access and accidental disclosure
- Ensuring that plans are in place to mitigate damage in the event of a security breach, unauthorized access or accidental disclosure

You should draft transparent data policies and privacy procedures for your project to allow individuals to:

- Understand and manage what data is being collected, how it is being used and disclosed, and how long it will be kept
- Give meaningful consent for automatic data collection, and for any after-the-fact data sharing
- Review, obtain and correct inaccuracies in their personal data

To ensure that privacy is considered at every stage of IRCC's advanced analytics projects, a privacy expert works within advanced analytics teams. Beyond supporting a more seamless process, this means that IRCC's privacy experts are better equipped to provide guidance related to the unique challenges and opportunities when it comes to AI and automation. The Government Advisory Directorate at the Office of the Privacy Commissioner (OPC) can also provide advice on privacy matters related to your project, even if you are not undertaking a Privacy Impact Assessment. To do so, start by contact the ATIP Division, IRCC's central point of contact for dealing with the OPC.

Add observations, analysis and links to key documents here.

Working in the open

Have you considered whether information of the work you are doing should be shared externally?

While the use of AI in decision-making is an emerging and rapidly evolving domain, chances are, other teams are thinking about similar issues and facing similar problems. While data, personal information and coding of a sensitive nature still needs to be protected, there may be other types of information related to your project that could be shared openly and reused by others.

If you have conducted research, prepared reports or developed guidance about AI and automation that could be released without compromising privacy, security or program integrity, consider whether it

could be made available through IRCC's website or another open source development platform, like GitHub or the Government's Open Data portal.

Add observations, analysis and links to key documents here.

Will your system use open source software?

The Policy on Service and Digital requires departments to use open standards and open source software by default, with exceptions for circumstances where an open source option is not available or does not meet user needs.

If not using open source, departments are expected to favour platform-agnostic commercial off-the-shelf (COTS) options over proprietary COTS, avoiding technology dependency and allowing for substitutability and interoperability.

If developing a custom-built application, by default any source code written by the government must be released in an open format via Government of Canada websites and services designated by TBS.

See requirements set out in the TBS *Directive on Automated Decision-Making* around access to components and release of source code, and also the guidance provided in the Guideline on Service and Digital.

Add observations, analysis and links to key documents here.

Accountability

Who will be the decision-maker of record?

IRPA Section 4.1 specifies that officers can make decisions, determinations or examinations using automated systems made available by the Minister. This does not mean that systems themselves become the decision-maker of record. Humans design system architectures, business rules, confidence thresholds and the like, and humans are ultimately accountable for every decision the Department makes. The Minister, in turn, is accountable to Parliament for all of the Department's activities.

You will need to determine who will approve the business rules that are used in your decision support system, and whether the same person will be logged as the decision-maker of record. If an automated system determines some, but not all, elements of an administrative decision, will the person making the final determination be considered the decision-maker of record for the entire decision? Or will certain sub-elements of the overall decision be treated as separate decisions with their own decision-maker of record (e.g. eligibility vs admissibility)?

Add observations, analysis and links to key documents here.

How will your system keep an audit trail for each decision?

To prepare for potential legal challenges, it is essential that IRCC establish a chain of evidence about how whole systems work, as well as a trail relating to each individual decision that goes through them.

- The automated system has the capacity to automatically generate a comprehensive audit trail of the administrative decision-making path, or you will retain an electronic record in the applicant's file
- All the key decision points are identifiable in audit trail
- All the key decision points within the automated system's logic are linked to the relevant legislation, policy or procedure
- All decisions are recorded and accessible by the system's user, a reviewer or an auditor
- The audit trail generated by the automated system can be used to help generate a notification of the decision (including a statement of reasons or other notification) where required
- The audit trail is secure from tampering (to provide protection and data integrity)
- The audit trail identifies precisely which version of an automated system was used for each decision it supports
- The audit trail includes a comprehensive and printable modification history, including: who created the record (with time and date); who has modified the record (with time and date); what was modified; for privacy and commercial-in-confidence matters, who has viewed the record (with time and date); who made the final decision (with time and date)
- The audit trail starts by identifying the authority or delegated authority identified in legislation
- The audit trail shows who the authorized decision-maker is
- The audit trail enables the recording of human intervention in automated processes, for example recording who is authorized to exercise intervention

For eTA and Visitor Record and Study Permit Extensions, the automation tab in GCMS allows users to clearly see all sub-activities, including when applications "passed" or were dropped out for manual review.

Add observations, analysis and links to key documents here.

Are you assembling documentation outlining the development and deployment of the system as a whole?

This should include design specifications, training data, system versions and any other information that could be valuable in meeting TBS Directive requirements or in the event of an audit, investigation, enforcement action or judicial proceeding.

Add observations, analysis and links to key documents here.

How will your approach to ongoing quality assurance ensure that your automated system consistently produces the right outputs?

You may decide to start by having humans review every decision that is put through the model, and then gradually scale back the share of decisions that are human-reviewed. You should also ensure that you test for unintended fettering of officers, bias and discrimination.

In the course of a quality assurance exercise, you should make sure to document your process and its results, providing enough detail to allow others to repeat the exercise or to review it if an audit or legal challenge occurs in the future.

Add observations, analysis and links to key documents here.

Security

Is your system exposed to an unacceptable level of cybersecurity risk?

Ensuring the security of environments where IRCC develops, tests and deploys AI models is critical to system integrity, compliance with privacy obligations, and public trust in the immigration system.

IRCC's IT Security Risk Management Framework aims to preserve the confidentiality, integrity and availability of departmental information and IT assets, protect the authenticity of electronic transactions, and ensure continuity of services and business operations. This framework fosters best practices, promotes the principles of continuous risk management throughout the system development lifecycle process, and lays a solid foundation for IRCC's Security Assessment and Authorization (SA&A) process.

The SA&A process (a requirement under the *Policy on Government Security*) provides a level of assurance based on the expected degree of injury that can reasonably be expected from unauthorized access and/or disclosure of data. Ultimately the SA&A provides a documented baseline of evidence of due diligence in the implementation of cybersecurity controls that supports the core principles of responsibility, privacy, chain of evidence, demonstrable transparency, and oversight.

This process should include:

- A system threat risk assessment
- A review of compliance with existing cybersecurity policies and identified security controls
- Confidence that the system – including the data it is using – is protected against hacking and manipulation

Add observations, analysis and links to key documents here.

C. Preparing for launch

This section will help to answer questions like: *Is IRCC truly ready to implement the new system? Have you engaged all the right people and is the Department prepared to respond to questions and unforeseen disruptions?*

The process for getting final approval to launch may depend on the scale of changes you are proposing.

Transparency

Have you finalized your approach to public transparency?

IRCC should disclose to clients and Canadians all instances (lines of business, specific functions) in which automated systems are used are used to support decision-making. IRCC privacy, legal, technical, and communications experts will need to be consulted to help determine what information will be disclosed publicly and how it will be disclosed.

Public disclosure about systems should include:

- A plain language description of the system, including examples
  - the purpose of the system
  - the rationale for its use
  - its role within the decision-making process
  - possible outcomes for people
- A general description of the data underpinning the system and where it came from (including a link to any anonymized training data that is publicly available)
- A general description of the types of data the model uses
- Performance information, such as accuracy/error rates, distribution of affected clients (e.g. by country, by program line), cost or productivity vs humans-only alternative

Fully disclosing source code to the public is highly transparent in spirit, but most people cannot make sense of code. Thus, what can be meaningfully revealed to the general public is different from what can be shared with a community of technical experts. Where legally permissible and feasible without exposing departmental systems to vulnerabilities, IRCC should disclose source code to a panel of technical experts, and at the same time provide a plain language explanation of the system's workings to the public at large. Adopting an "Open by Design" approach has the added benefit of safeguarding the Department from cyber-attacks or sophisticated fraud on the part of malefactors with technical skills, thereby balancing transparency with program integrity. The expert panel reviewing the source code and broader system design should produce a report that includes a section suitable for public consumption.

If there are any system details that you feel should be protected from public disclosure, you will need to confirm whether this information can be protected. For example, are there suitable exemptions under the *Access to Information Act* and other legal protections available, such as for protection of national security (and, if applicable, under the TBS Directive)?

Add observations, analysis and links to key documents here.

Have you prepared a notice for clients, alerting them to the fact that a decision on their application will be made in whole or in part by an automated system?

Notices at the point of collection of personal information should be written in plain language and easily findable on the program webpage. The TBS *Directive on Privacy Practices* and the *Directive on Automated Decision-Making* provide additional detail about the requirements for privacy notices.

The following except in the Privacy Notice Statement that appears on the TRV application form:

> *The personal information collected on an application, and other information collected in support of an application, may be used for computer analytics to support processing of applications and decision making, including your application. Personal information, including from computer analytics, may also be used for purposes including research, statistics, program and policy evaluation, internal audit, compliance, risk management, strategy development and reporting.*

For guidance around privacy notice statements, please contact the ATIP Division.

Add observations, analysis and links to key documents here.

Have you consulted your communications advisor?

- Have you identified necessary updates to IRCC's external-facing page on automated decision-making?
- Will a communications strategy needed?
- Will there be an announcement?
- Have you prepared key messages, Qs&As, etc.?

Add observations, analysis and links to key documents here.

Accountability

Have you worked out a governance structure that sets clear roles and responsibilities for the main partners delivering the project, including when it transitions from pilot to steady state?

The TBS *Directive on Automated Decision-Making* specifies that the Assistant Deputy Minister responsible for the program using an automated decision system is responsible for requirements related to algorithmic impact assessments, transparency measures, quality assurance, client recourse, and reporting. However, specific roles and responsibilities will vary according to IRCC's broader governance structures and individual projects.

Add observations, analysis and links to key documents here.

Have you met all privacy requirements and addressed any risks that were identified in your Privacy Assessment or by the Office of the Privacy Commissioner (via your engagement with IRCC's ATIP Division)?

Engage the ATIP Division early on to make sure that measures to mitigate privacy risks can be built in during the design of your initiative, rather than at the end.

Add observations, analysis and links to key documents here.

Have you met all Security Assessment and Authorization requirements and addressed any risks that were identified in your Security Action Plan (SAP)? Do you have an Authorization Decision Letter?

Add observations, analysis and links to key documents here.

Do you have a strategy to ensure that project managers can track progress, measure and report on results, and support future audits or evaluations?

Ensure that you have established appropriate metrics to measure specifically the impact of the automated function(s) of your program.

Add observations, analysis and links to key documents here.

Peer review

Have you consulted appropriate experts who can review the design of your system?

Depending on the potential impact of your project, the *TBS Directive on Automated Decision-Making* may require you to do this. Even if not compulsory, consultation with experts in AI or data science (e.g. from the National Research Council or Statistics Canada) may yield helpful suggestions for an improved system.

Add observations, analysis and links to key documents here.

User readiness

Have you provided officers with training and instructions that are clear and user-friendly?

Understanding the strengths and limitations of the system will help users comprehend how it can best support them and how to use it appropriately. If officers don't have a high level of confidence in how an automated system works, they may be more likely to disregard the insights it generates. On the other hand, if officers place too much confidence in the system or are unsure about how to interpret results, they may tend to defer to the system, which could be seen as fettering discretion.

— Develop training materials and ask users to vet content for clarity and relevance.

- Make FAQ documents readily available to officers and other staff.
- Working with appropriate experts, consider whether Program Delivery Instructions need to be updated to explain the roles and responsibilities of employees and their use of information coming from these systems.
- Establish feedback loops to allow officers to provide input on their use of the system(s). This type of mechanism will ensure that users can share their first-hand experience and make their voices heard when challenges arise.
- Develop change management strategies, as needed. It is important not to underestimate the impact that automated systems can have on individuals, teams and our organization as a whole. The Culture and Change Management Branch can support you in identifying appropriate measures to support a smooth transition.

Add observations, analysis and links to key documents here.

System readiness

Have you devised a contingency plan in the event of a system outage (short-term and long-term)?

See TBS requirements related to contingency systems and processes in the *Directive on Automated Decision-Making.*

Add observations, analysis and links to key documents here.

Partner readiness

Have you confirmed with affected partners – both internal and external – that they are prepared for the launch of your system?

Consider whether partners will require any support, such as briefing materials or communications products.

Add observations, analysis and links to key documents here.

Recourse

Are there sufficient recourse avenues for clients who wish to challenge the decision on their application? Does the introduction of automation necessitate any changes or additional pathways?

When a client challenges a decision that was made with the support of automation, and IRCC is authorized and agrees to review it, the review process should be such that (A) the decision reviewer stands in the shoes of the original decision-maker, and (B) the reviewer is not bound by the original decision, nor are they bound to take into account any findings or recommendations offered by the computer system that the original decision-maker relied upon. If the same technology used by the

original decision-maker is also used by the decision reviewer, there is a risk that a flaw or error in the technology will carry over from the initial decision to the review.

Add observations, analysis and links to key documents here.

D. Once up and running

This section will help to answer questions like: *Is the system still functioning as originally intended? Do any intervening factors point to the need for a review?*

Quality assurance and monitoring

Are you conducting ongoing quality assurance to verify that your system is not generating errors or unintentional outcomes?

For example, this could include reserving a significant share of overall applications for human-only review or randomly selecting cases that go through the automated system and mixing them in with human decision-makers' normal caseload. In both cases, you would need to make sure that officers don't know which applications are new and which are quality assurance checks.

You should also monitor the operation of the system to detect and avoid administrative law and privacy risks, and determine whether any residual discriminatory impacts can be detected.

Add observations, analysis and links to key documents here.

Do your results to date suggest that IRCC could safely lower the system's confidence threshold in order to realize greater efficiencies? Or the reverse?

Internal stakeholders, including Legal Services and relevant program policy areas, should be involved in reviewing results and changes to model rules.

Add observations, analysis and links to key documents here.

Are you monitoring for environmental security?

This should include monitoring for new cybersecurity threats and overall system integrity.

Add observations, analysis and links to key documents here.

Are you monitoring the results against the anticipated impacts you identified in previous stages of the project – e.g. in your GBA+ analysis or diagnostics?

This approach will allow you to assess whether measures to mitigate risk to diverse populations are adequate, and can help identify any additional unintended negative consequences.

Add observations, analysis and links to key documents here.

Has any new data become available that could improve the next iteration of your decision support system?

If so, consider whether new data could replace proxies with more direct evidence of outcomes.

Add observations, analysis and links to key documents here.

Has your model generated new insights about your program?

For instance, the expression of interest system facilitated by Express Entry has highlighted opportunities to modernize the entire federal economic program. As well, the data analytics underpinning the China TRV predictive model has bolstered the rationale for the CAN+ program (i.e. travel history).

You will need to consider questions like whether the new system will effectively replace program requirements established in legislation or regulations with algorithmically-generated business rules.

Add observations, analysis and links to key documents here.

Is there an established cycle and process for periodic review of the system's business rules, data quality, confidence thresholds and productivity gains?

Without an established approach to periodic review, there is a risk that the proper functioning of a system will depend on certain key individuals staying in their same job. IRCC needs to identify limitations of algorithms (and the data underpinning them) consistently to decision-makers, regardless of staff turnover or changing roles.

IT Security Authorization Letters must be renewed at least once every 5 years, or upon significant changes to the system.

Add observations, analysis and links to key documents here.

Are you assessing the overall performance of the system and reporting on results?

Depending on the objectives you established at the start, you may wish to consider how automation has helped us to reduce costs, save time, detect risks to program integrity, or improve client satisfaction.

You will also need to assess and report on how the system is contributing to meeting overall program objectives (see reporting requirements set out in the *Directive on Automated Decision-Making).*

Add observations, analysis and links to key documents here.

User insights

Are you monitoring the use of automated systems and gathering feedback from users?
As human responses to an automated system can vary over time (and the humans themselves can change as a result of staff turnover), systems should be monitored on an ongoing basis to avoid unintended negative outcomes.

- Involve users in ongoing quality assurance processes. A portion of client applications should be reserved for manual processing by officers. In addition to controlling for consistency between officer and machine decisions (e.g. through "blind testing"), this will help officers maintain the skills necessary to process applications with a range of levels of complexity.
- Leverage officers' knowledge to inform system changes and updates. Taking into account human intelligence will ensure that the system remains relevant in light of evolving country-specific risks and trends. Before scaling up or transitioning from a pilot phase to a steady state, engage a representative group of users to gather their feedback, observations and lessons learned.
- Monitor changes in users' behaviour and use of the system. You designed the system with the user in mind, you tested and tweaked machine outputs and interface, and you communicated relevant information to officers. However, it is possible that behaviour will change over time. As officers gain experience interacting with the system, they may become more, or less, trusting of machine outputs. To detect inadvertent automation bias or overreliance on system outputs, consider engaging users for periodic focus groups, observing users' behaviour in real environments, and analyzing data for anomalies. Users should also be involved any time you are considering changes that could impact the design of system outputs or interface.
- Keep users in the loop by circulating reports on the system's performance. This will demonstrate openness and transparency, and provide users with greater appreciation of the benefits (and challenges) related to automated support for decision-making.

Add observations, analysis and links to key documents here.

Transitioning from pilot to steady state

Taking your system from the pilot phase to mainstream implementation is a pivotal moment to take stock. Work with a multidisciplinary team to review questions like what worked well, what didn't, and what you would change if you make if only you knew then what you know now.

- Have you identified transitioning to steady state require additional safeguards, engagement or disclosure?
- Will your project require additional software or hardware?
- Have you tested the project for scalability?
- Have you learned lessons over the course of the pilot that could inform IRCC's future use of automated support for decision-making?

Establishing a formal reporting at the end of a pilot is not only good practice for effective project management. It also strengthens accountability and transparency, and provides greater opportunity for others to learn from your experience.

Add observations, analysis and links to key documents here.

PART III – AN OVERVIEW OF LEGAL CONSIDERATIONS AND PRACTICAL TIPS

The rule of law helps to ensure that the law is administered in a transparent and predictable manner, and provides a form of accountability to individuals affected by administrative decisions. The same legal principles and obligations of public law apply when AI is used to support administrative decision-making.

Highlighted below are some key legal considerations that arise – specifically with respect to administrative, human rights and privacy law – when using AI and automation to support administrative decision-making. These are intended to guide IRCC staff throughout the AI lifecycle, and include a set of practical tips to help mitigate risk.

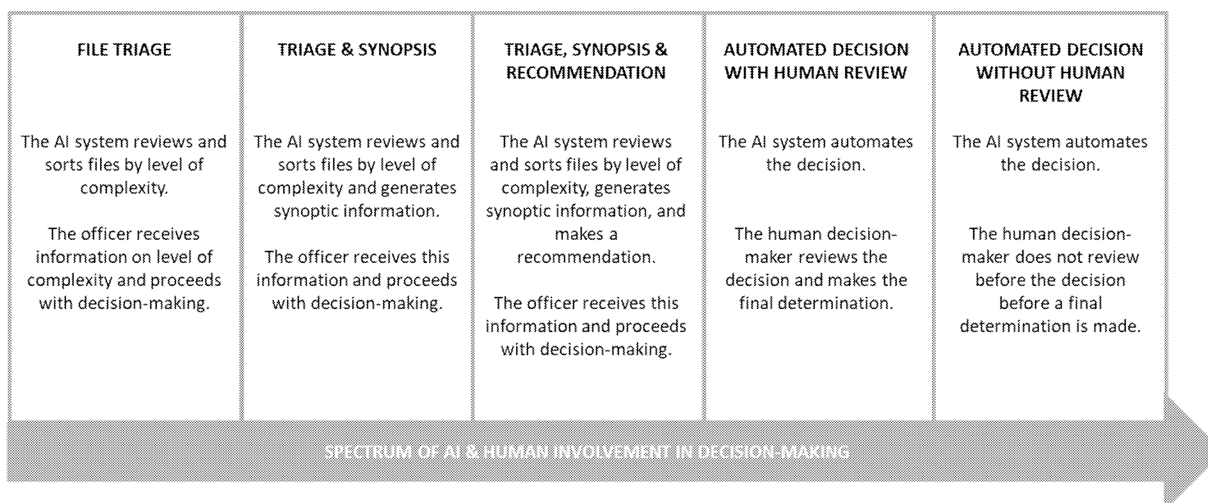
Working with a variety of multidisciplinary experts – including IRCC's policy and programs, operations, data, privacy, risk, client experience, communications, and transformation branches – will also help you to identify legal risk and adopt relevant mitigation strategies throughout various stages of your project. Given that legal considerations are context specific, **you must also engage IRCC's Legal Services Unit on an ongoing basis** to seek advice pertaining to the use case at hand (this is also required under the TBS *Directive on Automated Decision-Making*).

Administrative law

Administrative law accountability focuses on three essential concepts: legality, fairness, and reasonableness of decision-making. Issues that require particular attention in IRCC's environment include:

- Ensuring procedural fairness – an essential element of procedural fairness is meaningful participation. This means that the legislative criteria and the factual information used for decision-making are known to the applicant in advance, and that the individual has an opportunity to respond to it;
- Ensuring defensible decision-making – decisions are reasonable with a sufficiently transparent and intelligible justification;
- Avoiding undue or inappropriate reliance on AI – officers understand the intended scope of purpose of an AI system, and use the information provided by the system accordingly; and
- Considering the appropriate criteria and information, and mitigating bias – the information used by an AI system is consistent with legislative criteria.

A number of factors, including the degree of human involvement in design and development of AI systems and the basis for system changes over time, may impact administrative laws considerations. However, in general, the more an AI tool affects the decision-making process, the more care must be taken to ensuring that the principles of administrative law continue to be respected. A similar sliding scale of risk mitigation requirements calibrated to the level of impact of automated systems is reflected in the Government of Canada's Algorithmic Impact Assessment tool and *Directive on Automated Decision-Making*.



Human rights law

Human rights law provides another important lens for thinking through the legal considerations that arise when using AI to support decision-making. Human rights law is relevant both because of the nature of certain administrative decisions, and because the use of AI tools can give rise to human rights considerations even in situations where the administrative decisions being made would not ordinarily affect an individual's human rights. For example, an AI system may use a client's address or geographic location information inadvertently as a proxy for national or ethnic origin, or membership in a particular religious group. Human rights protections related to equality and discrimination are of particular relevance in IRCC's context.

Two key questions to ask when developing and deploying an AI model are:

1) **Does the model create distinctions based on enumerated or analogous grounds?** Enumerated grounds are those explicitly listed under Section 15 of the Canadian Charter of Rights and Freedoms, i.e. race, national or ethnic origin, colour, religion, sex, age, or mental or physical disability. Analogous grounds are prohibited grounds of discrimination identified by the courts, e.g. non-citizenship, marital status, sexual orientation, status of being adopted, and place of birth.

2) **Are these distinctions discriminatory?** Note that not all differential treatment is discriminatory. In some cases, it may be relevant to take into account personal characteristics in light of legislative criteria.

Privacy law

Under the *Privacy Act*, accountability for government institutions focuses on principles and obligations related to the collection, notice, use, accuracy, disclosure, retention, and disposal of personal information. While the Act is designed to be technologically neutral, there may be greater potential for privacy intrusion in the context of AI and automation given the sheer volume of personal information involved.

IRCC staff involved in AI projects must practice strong data and records management, and take steps to ensure that the use of personal information in AI tools is consistent with the purposes for which it was initially collected. Familiarize yourself with applicable privacy rules and requirements and consult with privacy experts at the outset to ensure privacy considerations remain front of mind during the entire AI lifecycle.

Practical tips to align with administrative, human rights and privacy law

The following practical tips suggest some steps to effectively 'put the law into practice' when approaching AI projects. They are designed to assist practitioners in moving from interpreting and applying the legal frameworks, to identifying key risks in the areas of administrative, human rights and privacy law, and mitigating these risks.

- **Consider what personal information may be used to accomplish an intended purpose, and use the information with the lowest privacy interest.** Records of model testing and any identified tradeoffs should be kept to document the rationale for decisions affecting the use of personal information.
- **Ensure notification obligations are met in relation to new collections of personal information or new uses of previously collected personal information.** IRCC can re-use personal information if the proposed use is consistent with the purpose for which it was originally collected, or with the consent of affected individuals. However, this requires careful review to establish a sufficiently direct connection between the initial purpose for which personal information was collected and the proposed purpose for an AI system.
- **Ensure data used is accurate, up to date, reliable and as complete as possible**. This should include assessing the suitability of training data for intended purposes, its reliability as a representative sample, and the accuracy of personal information in the data set over the lifecycle of its intended use. Working with IRCC's Chief Data Officer Branch and relevant program and policy leads will help ensure that historical data is suitable for purposes such as training an AI model.
- **Assess the extent to which de-identification[2] can be used.** While effective in minimizing privacy intrusion, this technique can be challenging as linking information to an identifiable individual can occur as new information is collected and technical means to defeat de-identification become more sophisticated.
- **Evaluate data flows to ensure collection, use, and disclosures of personal information meets legal and policy requirements.** Model developers and data experts will ensure a strong understanding of how the model works and which data it uses, whereas privacy experts will be able to provide advice on whether or not you need to complete a privacy impact assessment.
- **Assess training data for bias and evaluate whether the model/rules can either detect and correct, or, potentially embed or exacerbate bias**. This should be done in the early stages of design and on an ongoing basis as bias can involuntarily creep into a model over time. AI can

[2] De-identification refers to the process of modifying data such that it can no longer be attributed to an identifiable individual and cannot be reengineered to re-identify an individual.

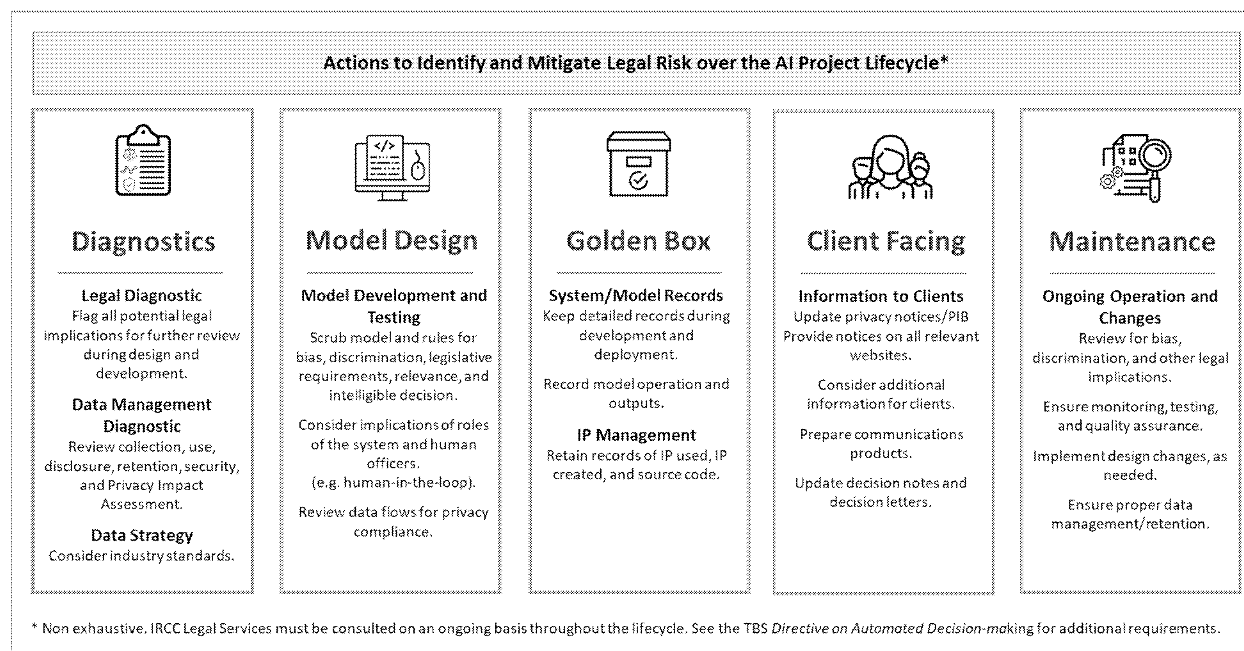
be leveraged to detect and mitigate individual human cognitive bias and discrimination, enhancing the quality and consistency of decision-making.

- **Consider functionality differences with AI models and whether they impact privacy obligations.** This could involve assessing the potential impact of a privacy breach for different models or options under consideration to see how it impacts overall risk.
- **Evaluate the *algorithm* and its suitability for the intended purpose including based on best practice, industry standards (if any), and any other applicable rules or regulations.** In addition to engaging internal partners, you may need to seek advice from external stakeholders through a formal or informal peer review process.
- **Evaluate the model and rules for inclusion and consideration of legislative criteria and other relevant information**. This should include review by an experienced officer to ensure that the model and rules are consistent with traditional decision-making processes.
- **Identify whether input variables or model rules take into account enumerated or analogous grounds directly, indirectly or by proxy.** This requires a detailed understanding of the data and context, and will require collaboration with model developers, data experts, and tech counsel.
- **Determine whether it is relevant and appropriate to consider enumerated or analogous grounds, for instance, in light of legislative criteria or its relevance to decision-making.** Tech counsel, policy/program leads and experience officers can provide expertise to make informed choices in this respect. For example, the legislative criteria for becoming an economic immigrant take into account the applicant's age, but when considered alongside other factors, age is a relevant indicator of things such as the length of time the applicant is likely to remain in the workforce and their ability to adapt to life in Canada.
- **Look for patterns that may suggest differential treatment based on enumerated or analogous grounds, and assess whether this could result in discriminatory outcomes.** This assessment should be made in consultation with IRCC's tech law team, experienced officers and other multidisciplinary experts. You may need to undertake further analysis to determine the root cause (e.g. training data, algorithm, human influence, rule design) and take steps to address the issue.
- **Assess the technology and its use by decision-makers with the help of multidisciplinary experts to meet transparency and explainability requirements.** Capturing the reasons for a decision may be done in one of two ways: creating and retaining of the decision on each application, or creating the ability to reconstruct processing steps and reasons for decision after the fact.
- **Assess intelligibility and cogency of decisions and reasons available.** Depending on the context, decisions within a range of acceptable outcomes are defensible – e.g. avoiding undue reliance on AI or fettering discretion, avoiding bias, applying the legislative criteria, appropriate consideration of relevant information, and cogent reasons. AI can also be used to help decision-makers provide reasons for their decisions that are more easily understood by affected individuals, increasing accountability and potentially facilitating access to judicial review.
- **Evaluate record keeping requirements.** Information used for administrative decision-making purposes must be retained for a minimum of two years. Consideration should also be given to

keeping records in relation to 'privacy by design' choices that were made in designing and deploying AI tools.
- **Provide adequate information and training to decision-makers** to ensure an understanding of the AI tool and the respective role of the AI tool and the human decision-maker.
- **Monitor the operation of the AI system on an ongoing basis** to detect and avoid administrative, human rights and privacy law risks. This may include macro-level analysis to detect any remaining discriminatory impacts, or monitoring of privacy or administrative law risks that may emerge over time.

Engaging in the AI Project Lifecycle

The figure below provides a high-level snapshot of steps to be taken over the AI project lifecycle – from developing diagnostics and appropriate model design early on to providing adequate notice to clients and ensuring ongoing monitoring. Records management and retention – i.e. 'the golden box' – are also critical to the responsible use of AI in all stages of the lifecycle. This may include records of consultations, model rules and model changes, statistical analysis, quality assurance, key decisions about design and deployment, instructions provided to officers, and steps taken mitigate legal risk on an ongoing basis.



PART IV – BASELINE PRIVACY REQUIREMENTS

In Canada, privacy is considered a human right. As the majority of the data we handle at IRCC is personal information, privacy requirements must be top-of-mind when planning, developing and monitoring any initiative involving data-driven technology. Privacy implications and nuances will change based on the specifics of the initiative, however there are baseline requirements for all initiatives that must be met.

This document is intended to aid programs and policy leads in considering privacy at the outset of a disruptive technology initiative and to help guide the implementation or modification of procedures, processes and governance as may be appropriate to ensure adherence to these baseline requirements.

Throughout this document, the term 'disruptive technology' is used to encompass data science techniques and data-driven technologies such as advanced analytics, machine learning, artificial intelligence, and automation. A disruptive technology is a new technology which has a rapid and major effect on technologies that existed before[3]. Additionally, this document is meant to be read in conjunction with the Policy Playbook on Automated Support for Decision-Making.

Within this document you will find privacy-specific requirements that are mandatory regardless of the disruptive technology or automation being considered. This means that these are the **minimum requirements** and there is room for further privacy protections. The privacy requirements laid out in this document are based on the *Privacy Act*, Treasury Board Secretariat policies, directives, guidelines and internal IRCC guidance.

Following the requirements set out below is **not** a substitute for a formal Privacy Impact Assessment (PIA). Rather, these requirements will guide programs in figuring out what they have to do to be privacy compliant and may help facilitate the completion of privacy assessments as may be required. Programs intending to add disruptive technology to their programs should fill out a Privacy Needs Assessment (internal link) and send it to the ATIP Division: ATIPinternal-AIPRPinterne@cic.gc.ca.

Legal Authority

A program must identify the parliamentary authority to collect and use personal information for the specified purposes of the program.

At IRCC, Part 4.1 of the Immigration Refugees Protection Act grants programs that fall under IRPA the ability to use electronic means to administer the Act, and section 2.2 of the Passport Order grants the Passport Program that authority. If you are unsure of your legal authority to collect and use personal information in a disruptive technology-enabled initiative, contact Legal Services.

[3] Adapted definition from Collins Dictionary: https://www.collinsdictionary.com/dictionary/english/disruptive-technology . Accessed on January 19, 2021.

Collection

> Personal information must only be collected if it relates directly to an operating program or activity of IRCC and each personal information data element must be necessary to the administration of the program[4]. When possible, personal information should be collected directly from the individual[5].

For the purposes of disruptive technology initiatives, **only information found in departmental systems of record (ex. GCMS) should be used**, and the activities of training models and algorithms should be done outside of those systems of record. Data collected from outside sources should not be used unless demonstrably necessary, and proper Information Sharing Agreements, Memoranda of Understanding, service level agreements etc. should be in place in relation to this use and strictly followed.

When considering using personal information in a model or algorithm, answer the following questions[6]:

- Demonstrating necessity: Is there a clearly defined necessity for the collection and use of the personal information? What is the problem that is expected to be solved by including this information?
- Proportionality: Is the collection and use of personal information targeted and tailored enough, so as to be viewed as reasonably proportionate to the reduction of privacy rights of the individual that the collection may represent?
- Effectiveness: Is this personal information effective at treating the issue, and so clearly connected to solving the problem?
- Minimal intrusiveness: Are the specific data elements chosen the least intrusive information available? Ensure that all other less intrusive avenues of investigation have been exhausted.
- Accuracy and validity: Is the data source a reliable and valid source of information? How will you ensure the data is accurate and up-to-date?

Notice and Informed of Purpose

> IRCC must notify individuals of the purpose for which their information is being collected, commonly referred to as a 'privacy notice'[7]. This notice must be given at or before the time of collection.

Depending on the technology being used, **it may be necessary to update existing privacy notices** and Personal Information Banks (PIB) to account for the use of the technology. If the personal information is to be used or disclosed for a purpose other than for the original purpose or a consistent purpose, a formal privacy assessment will likely be required.

---

[4] Section 4 of the *Privacy Act.*
[5] Sub-section 5(1) of the *Privacy Act.*
[6] Adapted from the Office of the Privacy Commissioner's publication *A Matter of Trust: Integrating Privacy and Public Safety in the 21st Century*: https://www.priv.gc.ca/en/privacy-topics/surveillance/police-and-public-safety/gd_sec_201011/#toc5a. Accessed January 29, 2021.
[7] Sub-section 5(2) of the *Privacy Act.*

Transparency

IRCC must notify past applicants that their information was used to train or build models.

Updating IRCC's public-facing website about analytics and automation is one suggested way of notifying the public that past applications have informed current models. An online notice that can be modified as IRCC modifies its use of technology may also help to achieve transparency requirements.

Explainability

Individuals have a right to know exactly how their personal information was processed through a disruptive technology system. Ensuring that plain language explanations are available on demand would allow individuals to see how technology was used to support decision-making.

Accuracy

IRCC must take all reasonable steps to ensure that personal information used for an administrative purpose[8] is as accurate, up-to-date and complete as possible[9]. This also includes ensuring there are mechanisms in place to correct inaccurate information.

For initiatives involving disruptive technology, this involves ensuring data is collected from a reliable source, the quality of the data, developing technological mechanisms to ensure the technology is working (such as feedback loops and blind tests), quality assurance on the outputs, and so on. Additionally, to ensure the accuracy of the data, program areas must take the necessary steps to minimize unintended bias in the data. Finally, accuracy also involves model maintenance and ensuring the model is trained and re-trained on the most updated, accurate and reliable data.

Use

Personal information must only be used for the purpose it was initially collected, a use consistent with that purpose or for a purpose for which it may be disclosed under section 8(2) (discussed under Disclosure, below)[10].

Applying disruptive technology to a dataset involving personal information is a use; this includes all uses whether administrative (as a part of a decision-making process that affects the individual) or not. Personal information must be treated appropriately regardless of the level of automation or support the technology is providing. You must also ensure that the use of disruptive technology is a consistent use of the personal information. To determine what constitutes a consistent use of personal information, the

[8] Annex A of the Interim Policy on Privacy Protection.
[9] Sub-section 6(2) of the *Privacy Act.*
[10] Section 7 of the *Privacy Act.*

original purpose and the proposed purpose must be so closely related that the individual would expect that the information would be used for the consistent purpose, even if the use is not spelled out. Programs that are unsure if it is a consistent use of the personal information should consult Legal Services.

The following can help to reduce the likelihood and impact of inappropriate use:

- **Data Minimization**: Only use the personal data elements that are absolutely necessary. For example: rather than full name and UCI, only the UCI is required, or rather than either, a unique identifying number can be generated for the purposes of the model only.
- **Reducing Data Granularity**: This involves removing the precision of some data elements where precision is not necessary. For example: the whole birth date might not be required, only the birth year would suffice and generate effective outputs; or the whole phone number might not be required, only the area code could suffice.
- **De-identification**: It may be possible to de-identify all or parts of datasets to reduce the likelihood of identifying individuals and reduce risks to individual privacy. Data sets may even be de-identified to the extent that it is no longer personal information which may lend to fewer data-related restrictions, yet still produce valuable results.
- **Need to Know Principle**: Only those individuals who need to know the information should have access to that information, regardless of the type of disruptive technology being applied. This may mean creating different versions of outputs: a detailed one for officers that includes a high degree of personal information, and a summary output for management that has less personal information. It may also mean limited individuals have access to the training data and where the outputs are saved. Saving information in ways that allow for modification and tailoring of access rights can help with this.

Disclosure

Personal information under IRCC's control must not be disclosed to anyone or any organization for any reason, except for those reasons listed in sub-section 8(2) of the *Privacy Act*[11].

For initiatives involving disruptive technology, this includes information found in departmental systems of record (ex. GCMS) that other organizations such as CBSA or CSIS can view. Regular information sharing may continue to occur between IRCC and partner organizations, however Memoranda of Understanding (MOUs), Information Sharing Agreements (ISAs) and other formal agreements must be updated and modified through the appropriate channels if there is the desire to disclose the outputs on a regular basis. IRCC's Admissibility and International and Intergovernmental Relations (internal link) groups can provide further guidance.

Safeguards

Personal information must be safeguarded appropriately regardless of the kind of technology applied to it.

[11] Section 8 of the *Privacy Act.*

Appropriate administrative, technical and physical safeguards should be applied to personal information for every program using personal information and therefore extends to at all stages of a disruptive technology initiative. Much like other program efforts and operations, consideration should be given to reducing the likelihood of privacy and security breaches throughout development of new technological solutions. Programs are responsible for completing the mandatory IT Security Assessment and Authorization process (internal link). Additionally, programs are responsible for developing measures to reduce the impact in the event of a privacy or security breach. IRCC's IT Security experts can offer more detailed guidance.

Retention and Disposal

Personal information used for an administrative purpose (as a part of a decision-making process that affects the individual) must be retained for at least two years, and in accordance with the appropriate Retention and Disposition Schedule[12]. All data (with the exception of training data in the Exploration Zone) must be kept such that in the event of a complaint or legal action, the decision can be replicated.

For disruptive technology initiatives, retention and disposal decisions must be made on:

- Training data: Should be destroyed or removed from the Exploration Zone when it is no longer required for training purposes;
- Model/algorithm: Should be versioned similarly to software retention practices; and
- Outputs: Depending on the kind of model, the outputs should be added to an existing Retention and Disposition Schedule, a new RDS be created, or the outputs be deemed transitory and destroyed when no longer required.

IRCC's Information Management Branch (internal link) can provide guidance on the retention and disposition of records created around the development and deployment of disruptive technology.

Individual Access

Individuals have a right to access the personal information that IRCC has about them[13]. IRCC should always have ways to disclose personal information on-demand.

Records around the development, deployment and monitoring of disruptive technology initiatives should also be kept at hand in the event of an *Access to Information Act* or a *Privacy Act* request.

[12] Sub-section 6(1) of the *Privacy Act.*
[13] Sub-section 12(1) of the *Privacy Act.*

Complaints and Reviews

Keep clear and retrievable documents that outline all of the actions and decisions taken to ensure privacy compliance. They will be required to be produced in the event of a complaint or review.

Individuals have a right to file a formal complaint to the Office of the Privacy Commissioner (OPC) of Canada about how IRCC handles and manages their personal information, this includes how the information is processed through technology[14]. The OPC may also initiate reviews of programs when they deem necessary, and they have powers to obtain any and all information relating to the personal information handling of the program, including details about models, algorithms and automation.

Monitoring

Monitoring for privacy compliance to the above-noted requirements should be built in from the model development phase and a monitoring schedule post-deployment should be followed that involves both the program owner(s) as well as the architects of the model.

In addition to ensuring that the disruptive technology is working properly, here is a list of non-exhaustive monitoring activities to plan for:

- **Collection**: Make sure no data from sources other than departmental systems of record (ex. GCMS) and other IRCC data repositories are included in the disruptive technology, and if there is outside data, ensure the appropriate ISA or MOU is in place and up to date.
- **Notice**: Review privacy notices and transparency and explainability practices to ensure accuracy and up to date information. Update whenever required.
- **Retention and Disposal**: Review the retention and disposal practices and ensure no information is retained for longer than it should be.
- **Accuracy**: Build in regular data quality practices to ensure data is accurate, up to date and complete as possible, and modify the information when required.
- **Use**: Review data handling practices, ensure mitigation measures against inappropriate use are functioning properly, employ new measures as required, and update practices periodically.
- **Disclosure**: Review disclosure practices to ensure disclosures are occurring as a part of up to date MOUs and/or ISAs, and ensure other government organizations can only see what information they are allowed to see in GCMS. Make modifications when required.
- **Safeguards**: Conduct regular security checks to ensure the training data, the technology itself and the outputs are secure.

This document provides an overview of the baseline privacy requirements that all disruptive technology initiatives at IRCC must consider. Additional requirements may be required based on specific initiatives, and even if you follow these requirements, a privacy assessment may be required. A multi-stakeholder approach is the best way to ensure all requirements and risks are addressed and mitigated

[14] Section 29 of the *Privacy Act.*

appropriately. The ATIP team and others throughout the department as referenced within this document are here to help you navigate the privacy considerations to ensure that programs are built with privacy in mind. Please fill out a Privacy Needs Assessment (internal link) to begin your consultation with the ATIP Division.

PART V – CHECKLIST FOR THE DIRECTIVE ON AUTOMATED DECISION-MAKING

The *Directive on Automate Decision-Making* is the Government of Canada's policy that governs the use of artificial intelligence to support administrative decisions. The Directive came into force in April 2020, and applies to automated systems, tools and models used by federal institutions to make or support decisions affecting external clients.

The following checklist provides an overview of the requirements you must address at various stages to comply with the Directive. You should think of these as minimum requirements; other requirements specific to IRCC's legal framework, policies and programs may also apply.

Some requirements apply to all systems that fall within the scope of the Directive, whereas others vary according to the level of impact associated with algorithm used. In addition, although the Directive only applies to automated systems in production (therefore excluding those in a testing environment), reviewing the requirements and mandatory Algorithmic Impact Assessment tool early on will help you anticipate and develop critical risk mitigation strategies in the design and testing phase of your project.

A. Exploring automated support for decision-making as a potential solution

- ☐ **Legal** – Have you consulted with IRCC's Legal Services? This must be done starting from the concept stage and continue on an ongoing basis.

B. Designing your system

- ☐ **Algorithm Impact Assessment (AIA)** – Have you completed a preliminary AIA and identified requirements related to the relevant the level of impact?
- ☐ **Access to components** – If applicable, have you determined the appropriate license for software components and, if using a proprietary license, addressed related requirements?
- ☐ **Data quality** – Have you validated that the data collected for, and used by, the automated decision system is relevant, accurate, up-to-date, and in accordance with the *Policy on Information Management* and the *Privacy Act*?
- ☐ **Security** – Have you assessed security risks related to the system and applied appropriate safeguards, as per the *Policy on Government Security*?
- ☐ **Human intervention** – Have you ensured that the system allows for human intervention? (*Levels III and IV only*)

C. Preparing for launch

- ☐ **AIA** – Have you finalized a full response to all questions in the AIA and taken steps to address all requirements relevant to the level of impact?
- ☐ **AIA** – Have you released the final version of the AIA in an accessible format?
- ☐ **Source code** – Have you released custom source code owned by the Government of Canada, as required? You will also need to determine the appropriate access restrictions to the released source code.

- ☐ **Testing and monitoring outcomes** – Have you developed processes to test data and information used by the automated decision system for unintended data biases and other factors that may result in unfair outcomes?
- ☐ **Testing and monitoring outcomes** – Have you developed processes to monitor outcomes on a scheduled basis, including monitoring for unintentional outcomes and verifying compliance with relevant legislation and this Directive?
- ☐ **Notice** – Have you provided notice on all service-delivery channels in use to inform applicants that the decision rendered will be undertaken in whole or in part by an automated system? NB: Notice must be provided prominently and in plain language; IRCC's Legal Services Unit, ATIP Division and Communications teams can provide advice on. (*Levels II-IV only*)
- ☐ **Peer review** – Have you consulted the appropriate qualified experts to review the system? (*Levels II-IV only*)
- ☐ **Employee training** – Have you developed adequate training for employees and put in place measures/a strategy to ensure employees receive the training? (*Levels II-IV only*)
- ☐ **Contingency** – Have you established contingency systems and/or processes? (*Levels III and IV only*)
- ☐ **Approval** – Have you obtained the appropriate level of approvals for the system to operate? (*This requirement applies largely to Level III and IV systems, however, the Assistant Deputy Minister responsible for the program using the automated system must ensure that all relevant requirements set out in the Directive are met.*)

D. Once up and running

- ☐ **AIA** – Have you updated the AIA to reflect changes in system functionality or scope? In the event of a change in level of impact, you may need to address additional requirements.
- ☐ **Explanations** – Have you provided a meaningful explanation to affected individuals of how and why the decision was made?
- ☐ **Results** – Are you storing the results of decisions, as required, and complying with monitoring and reporting requirements?
- ☐ **Testing and monitoring outcomes** – Is data used by the system routinely tested to ensure that it is still relevant, accurate, and up-to-date?
- ☐ **Recourse** – Have you provided clients with appropriate recourse options to challenge the automated administrative decision?
- ☐ **Reporting** – Have you published information about the effectiveness and efficiency of the automated decision system in meeting program objectives?

GLOSSARY OF KEY TERMS

Administrative decision-making: Decisions made by authorized officials of an institution (such as a board, tribunal, commission, government department, agency or Minister), where the outcome affects legal rights, privileges or interests. Central to the immigration system, administrative decision-making is also critical in other fields like international trade, taxation, broadcasting and transport.

Algorithm: A sequence of instructions, rules, and calculations executed by a computer in a particular order to yield a result, typically an answer to a specified problem. Multiple algorithms can be used in combination to solve complex problems.

Artificial intelligence: Encompassing a broad range of technologies and approaches, AI is essentially the field of computer science dedicated to solving cognitive problems commonly associated with human intelligence, such as learning, problem solving, and pattern recognition.

Automated decision support system: Includes any information technology designed to directly support a human decision-maker on an administrative decision (for example, by providing a recommendation), and/or designed to make an administrative decision in lieu of a human decision-maker. This includes systems like eTA or Visitor Record and Study Permit Extension automation in GCMS.

Big data: A popular colloquial term used to describe immense information assets that inundate organizations on a day-to-day basis. This data is often highly varied and needs to be processed through computer systems at a fast pace, making it impossible to effectively manage and yield business value without specific software tools and analytical methods.

Black Box: Opaque software tools working outside the scope of meaningful scrutiny and accountability. Usually deep learning systems. Their behaviour can be difficult to interpret and explain, raising concerns over explainability, transparency, and human control.

Business rules: A series of instructions (if-then propositions), from which an automated system is unable to deviate. Rule sets can be human-crafted (based on experience, common sense, research) or human-curated (proposed by advanced analytics models but reviewed and approved by people). Automated systems comprise business rules that cannot be modified without human intervention. This is distinct from *autonomous* systems, which have pre-determined goals but no pre-determined approaches to how they are achieved.

Data analytics (including advanced analytics): The pursuit of extracting meaning from large volumes of (sometimes high-velocity) data using specialized computer systems. These systems transform, organize, and model the data to draw conclusions and identify patterns.

Data governance: Efforts at a corporate level to formally manage data assets (e.g. who does what, when, and under what authority). Effective data governance ensures that data is of high-quality as it is created, extracted and manipulated for reporting and decision-making. A bridge between the business and IT spheres, data governance supports process innovation and improved client service.

Exploration zone: The exploration zone – also referred to as a "sandbox" – is the environment used for research, experimentation and testing related to advanced analytics and AI. Data, codes and software are isolated from those in production so that they can be tested securely.

"Fettering" of a decision-maker's discretion: Fettering occurs when a decision-maker does not genuinely exercise independent judgment in a matter. This can occur when a decision-maker binds him/herself to a fixed rule of policy, another person's opinion, or the outputs of a decision support system. Although an administrative decision-maker may properly be influenced by policy considerations and other factors, he or she must put his or her mind to the specific circumstances of the case and not focus blindly on one input (e.g. a risk score provided by an algorithmic system) to the exclusion of other relevant factors.

Procedural fairness: Crucial to administrative decision-making, procedural fairness requires that applicants:

— have meaningful participation (they know the case to be met and have an opportunity to meet it) and are provided with a fair and unbiased assessment of their application
— are informed of negative evidence and the decision-maker's concerns
— have a meaningful opportunity to provide a response to such evidence and concerns about their application
— are provided reasons for the decision

There are numerous factors that determine what, specifically, should be done to address these principles. One such factor is the importance of the decision from the applicant's perspective.

Process automation: Also called "business automation" (and sometimes even "digital transformation"), process automation is the use of digital technology to perform routine business processes in a workflow. Process automation can streamline a business for simplicity and improve productivity by taking mundane repetitive tasks from humans and giving them to machines that can do them faster. A wide variety of activities can be automated, or more often, partially automated, with human intervention maintained at strategic points within workflows. In the domain of administrative decision-making at IRCC, "process automation" is used in contrast with "automated decision support," the former referring to straightforward administrative tasks and the latter reserved for activities involving some degree of judgment.

Machine learning: A sub-category of artificial intelligence, machine learning refers to algorithms and statistical models that learn and improve from examples, data, and experience, rather than following pre-programmed rules. Machine learning systems effectively perform a specific task without using explicit instructions, relying on models and inference instead.

Source code: Computer program in its original programming language, human readable, before translation into object code. It consists of algorithms and computer instructions, and may include developer's comments.